**General Disclaimer**

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.
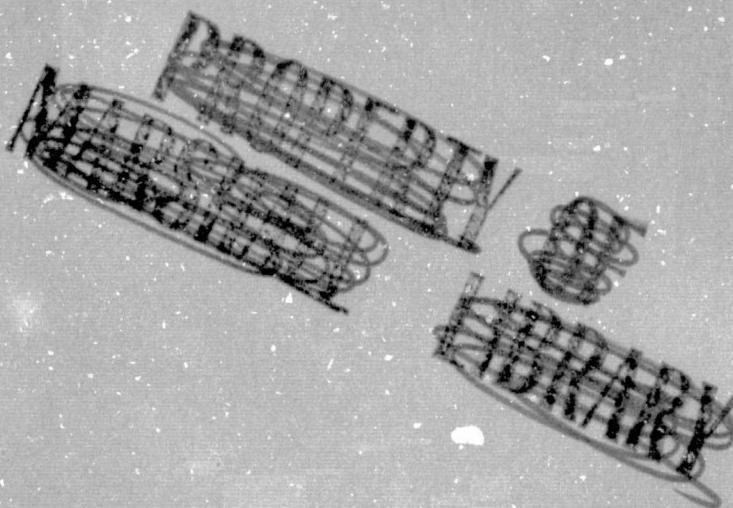
UCLA-ENG-7630
MARCH 1976

CONTROL BY MODEL ERROR ESTIMATION

R.E. SKELTON

Principal Investigator: P.W. LIKINS

UCLA • SCHOOL OF ENGINEERING AND APPLIED SCIENCE
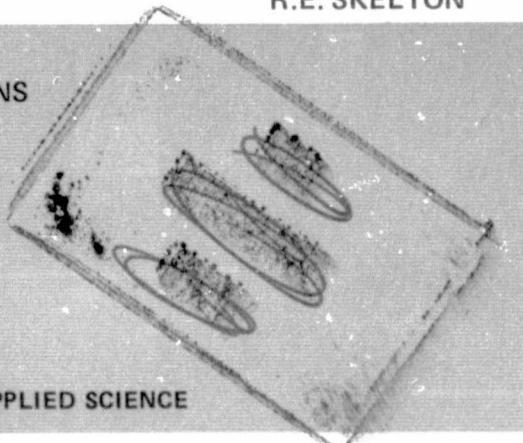
CONTROL BY MODEL ERROR ESTIMATION

Robert Eugene Skelton

Principal Investigator: Peter W. Likins

School of Engineering and Applied Science
University of California
Los Angeles, California

ABSTRACT

One of the most critical deficiencies of modern control

theory is the reliance of the theory upon the fidelity of the

mathematical model of the system. This research approaches the

problem of model errors in linear system models by augmenting to

the original system of equations an "error system"

$$\hat{e} = P\gamma$$
$$\dot{\gamma} = D\gamma \tag{1}$$

which is designed to approximate the effects of such model errors

as external disturbances, truncated modes, and parameter errors,

when this "error vector" $\hat{e}$ is added to the system equations to

obtain

$$\dot{x} = Ax + Bu + \hat{e} \tag{2}$$

Several researchers have concluded that increasing the order of the

model (via augmented "disturbance models," "dynamic compensators,"

etc.) can compensate for parameter errors,[1] and external

disturbances.[2-6] There is very little guidance available,

however, for the task of actually constructing such an auxiliary

system for augmentation, or for determining whether such methods

can also accommodate truncated modes, a problem which has in the

past been treated by singular perturbation.[7-9]

The point of view taken in this research is that the total of

truncated modes, external disturbances, and parameter errors can

be corrected by the addition of a "model error vector," $\hat{e}$, to the

original system equations. When state estimation techniques are

applied to the assumed "error system" (1), and when the error

system (1) belongs to a class called the Sturm-Liouville equations

(which generate orthogonal functions), the resulting estimator is

labeled an "orthogonal filter." A special case of the Sturm-

Liouville equations called the Chebyshev error system generates the

Chebyshev polynomials, which provide a least squares fit to the

measurement residual over an interval τ. The important selection

of this "observation window", τ, is related to a property of the

closed loop system called the characteristic time ($\tau \triangleq \max\limits_{x} \dfrac{V(x)}{-\dot{v}(x)}$,

where $v(x)$ is a Liapunov function for the asymptotically

stable linear system). Thus the parameter τ is both a model param-

eter and an element of the control system serving to interweave the

processes of model development and control system design.

In this dissertation a Chebyshev error system is developed for

application to the Large Space Telescope (LST), an earth satellite

which NASA plans to launch in the 1980's.

CONTENTS

FIGURES

SYMBOLS

$\mathscr{A}_0$ designation for a physical system

$\mathscr{A}_j$ designation for the jth model of a physical system

u^o control inputs to physical system

z^o actual measurements obtained from physical system

y^o actual outputs from physical system

x^j state vector of model $\mathscr{A}_j$; $x^j \in R^{n_j}$

y^j output vector of model $\mathscr{A}_j$; $y^j \in R^k$

z^j measurement of model $\mathscr{A}_j$; $z^j \in R^\ell$

$\tilde{y}^j$ $\triangleq y^o - y^j$ output residual

$\tilde{z}^j$ $\triangleq z^o - z^j$ measurement residual

Y $\triangleq \begin{pmatrix} y \\ z \end{pmatrix}$

$\tilde{y}^{ij}$ $\triangleq y^j - y^i$

e^{ij} the model error vector which causes $\tilde{y}^{ij} \equiv 0$

$\hat{e}^{ij}$ an approximation of e^{ij}

$\tilde{e}^{ij}$ $\triangleq e^{ij} - \hat{e}^{ij}$

$\left. \begin{aligned} A^j \\ \\ B^j \\ \\ C^j \\ \\ M^j \end{aligned} \right\}$ Parameters associated with the model $\mathscr{A}_j$

$$\dot{x}^j = A^j x^j + B^j u^o \; , \quad y^j = \begin{bmatrix} C^j \\ M^j \end{bmatrix} x^j$$

Q weighting matrix for output, y^j, in linear regulator problem

R weighting matrix for control, u^o, in linear regulator problem

K^j Riccati matrix associated with model $\mathscr{A}_j$

H weighting matrix for terminal $y^j(T)$ in linear regulator problem

$\hat{K}$ Kalman gain in Kalman filter (also more general estimator gain of full order estimator)

$\hat{x}^j$ estimate of x^j

G gain matrix used for the control law $u^o = G \hat{x}^j$.

$w(t)$ white noise random vector; $w(t) \in R^n$ ("plant noise")

$v(t)$ white noise random vector; $v(t) \in R^\ell$ ("measurement noise")

$\hat{Q}$ covariance matrix $E[w(t) \, w(t)^T] \triangleq \hat{Q}(t)$

$\hat{R}$ covariance matrix $E[v(t) \, v(t)^T] \triangleq \hat{R}(t)$

$\sum$ covariance matrix $E[x(t) \, x(t)^T] = \sum(t)$

s a sensitivity vector $s^T \triangleq \dfrac{\partial x}{\partial p}$

p a vector of parameters; $p \in R^r$

P a matrix satisfying $e^{31} = P\gamma$

D a matrix satisfying $\dot{\gamma} = D\gamma$

γ a vector of synthetic variables, characterizing the error system

τ an "observation window" for the state estimator

d the order of the model error system

d_{22}

d_{21}

B

T_1

T_2

$\hat{z}^t$

$\hat{x}$

Parameters of the state estimator

$$\dot{\hat{z}}^t = d_{22}\,\hat{z}^t + d_{21}\,z^o + B\,u^o$$

$$\hat{x} = T_1\,\hat{z}^t + T_2\,z^o$$

L

F

K

Partitioned parts of the Riccati matrix

$$K^2 \triangleq \begin{bmatrix} K & L \\ L^T & F \end{bmatrix}$$

$\lambda_i[D]$ i^{th} eigenvalue of the matrix D

$r[E]$ spectral radius of matrix E

$\overline{A}^j$ closed loop matrix; $\overline{A}^j \triangleq A^j - B^j\,R^{-1}\,{B^j}^T\,K^j$

T a transformation matrix such that

$$T^{-1}\,A^1\,T = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix}$$

$$T^{-1}\,B^j = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

$$C^j\,T = [C_1\ C_2]$$

$$M^j\,T = [M_1\ M_2]\ ,\ M^j = [\Gamma_{11}\ \Gamma_{12}]$$

$$T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}\ ,\ T^{-1} = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{bmatrix}$$

Γ $\triangleq [\Gamma_{21}\ \Gamma_{22}]$

$\cancel{x}$ a vector defined by $\cancel{x} = \begin{pmatrix} x^j \\ x^t \end{pmatrix}$

L_1 $\triangleq T_1$

L_2 $\triangleq T_2 \, T_{22}^{-1}$

T_1 $\triangleq \begin{bmatrix} T_{11} \\ T_{21} \end{bmatrix}$

T_2 $\triangleq \begin{bmatrix} T_{12} \\ T_{22} \end{bmatrix}$

$[\cdot]_{a,b}$ a matrix $a \times b$

$[\cdot]_a$ a square matrix $a \times a$

F_1 $\triangleq [0_{q\ell} \; , \; I_q] \, A^j \, L_1$

F_2 $\triangleq [0_{q\ell} \; , \; I_q] \, A^j \, L_2$

G_1 $\triangleq - M^j \, A^j \, L_1$

G_2 $\triangleq - M^j \, A^j \, L_2$

I_i an identity matrix $i \times i$.

σ normalized time $\sigma \triangleq \dfrac{2}{\tau} t - 1$

(See appendices for additional symbols)

1.0 Introduction

The task of determining a control policy so that the physical

dynamical system behaves in a desired manner usually begins with the

responsibility of obtaining a mathematical characterization (model) of

the physical system. However, any mathematical description of the phys-

ical system is subject to "model errors" in the sense that the mathe-

matical description does not precisely describe the dynamical behavior

of the physical system. Moreover, there presently is no well developed

theory to assist one in selecting a mathematical model of the physical

system which is appropriate for the particular control task at hand.

It is for this reason that model errors are a common cause of diver-

gence in the state estimators used in the feedback control policies.*

For linear models of systems and under other appropriate conditions,

such state estimators are commonly called "Kalman filters" in stochastic

characterizations of problems and "Luenberger observers" in deter-

ministic characterizations of problems. It is understood by users of

the theory that model errors establish performance limitations of the

closed loop physical system, and thus the realized performance may be

unacceptable even though the design is based upon optimal control theory.

The reason this model error problem persists in spite of the vol-

ume of "model reduction" literature that exists (see for example[10-12],

lies in the fact that the appropriateness of a model depends upon the

---

*Model errors also limit performance of classical designs. The modern
"state estimator" approach is utilized throughout this paper because
its structure includes classical designs but is not limited to them.
That is, a classical compensation design may be viewed a state esti-
mator for some model of the system.

1

nature of its (desired or undesired) inputs. As a consequence of this fact, the "modeling problem" and the "control problem" should not be treated as separate and independent problems. Efforts to treat them separately usually result in control policies which either

(a) suffer from modeling errors, causing performance requirements to be violated

or

(b) satisfy performance requirements at a high cost of control policy synthesis.

The conditions which cause (a) are sometimes not discovered before "flight" of the system. The condition (b) often prevents use of modern control theory in an application. To more sharply focus questions of model errors, attention is restricted to linearized models of the physical system.

In the quest for an appropriate system model (call it $\mathscr{A}_j$) it is necessary to consider the impact of the choice of models on the total closed loop system performance. It is helpful to consider the exist-ence of a hierarchy of linear models as shown in Figure 1-1. In this hierarchy, it is assumed that the physical quantity to be manipulated for control purposes has been selected (corresponding to its mathe-matical symbol, $u^o(t)$). Likewise, the mathematical symbol selected for the measurement vector is $z^o(t)$, corresponding to the selected physical quantities. The output vector, $y^o(t)$, represents those variables we actually wish to control, as opposed to the variables which are actually measured, $z^o(t)$. To complete the model, then, we must write some dif-ferential equations which dynamically connect the systems input

PHYSICAL SYSTEM

CONTROLS $u^o(t)$
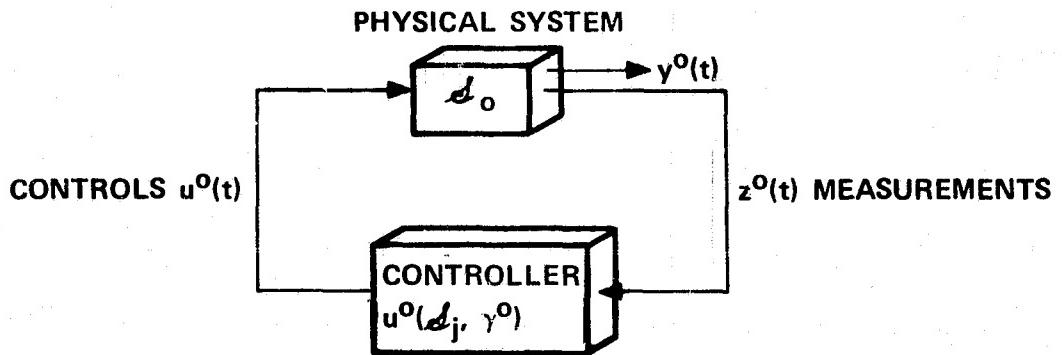
$z^o(t)$ MEASUREMENTS

$\mathscr{A}_o$: physical system

actual input $\quad u^o \in R^m$

actual output $\quad y^o \in R^k$

actual measurements $z^o \in R^\ell$

$$\mathscr{A}_1 \begin{cases} \dot{x}^1 = A^1 x^1 + B^1 u^o \\ \begin{pmatrix} y^1 \\ z^1 \end{pmatrix} = \begin{bmatrix} C^1 \\ M^1 \end{bmatrix} x^1 \end{cases}$$

$y^1 \in R^k$

$z^1 \in R^\ell$

$x^1 \in R^{n_1}$

$$\mathscr{A}_2 \begin{cases} \dot{x}^2 = A^2 x^2 + B^2 u^o \\ \begin{pmatrix} y^2 \\ z^2 \end{pmatrix} = \begin{bmatrix} C^2 \\ M^2 \end{bmatrix} x^2 \end{cases}$$

$y^2 \in R^k$

$z^2 \in R^\ell$

$x^2 \in R^{n_2}$

MODEL FIDELITY

$$\mathscr{A}_j \begin{cases} \dot{x}^j = A^j x^j + B^j u^o \\ \begin{pmatrix} y^j \\ z^j \end{pmatrix} = \begin{bmatrix} C^j \\ M^j \end{bmatrix} x^j \end{cases}$$

$y^j \in R^k$

$z^j \in R^\ell$

$x^j \in R^{n_j}$

Figure 1-1. Hierarchy of Linear Models.

(desired and undesired) to the system outputs. The order of these equations (or, the number of equations in the state form of Figure 1-1) remains to be determined; thus the hierarchy of possibilities illustrated in Figure 1-1.

The dilemma of model selection is complicated by two conflicting goals. We wish to implement the simplest solution to the control problem, for economic reasons, and we wish to achieve satisfactory performance. If the models of Figure 1-1 are arranged in order of decreasing model fidelity* (fidelity being defined according to their ability to faithfully predict the output of the physical system, $y^o(t)$ for a specified input and initial conditions), then the better closed loop system performance is achieved by designing the control policy on the basis of a model at the upper end of this model hierarchy, (since more information about the physical system is made available) whereas the simplest control policy obtains from the lower end of the model hierarchy. The limits of this argument lead us to the conclusion that truly optimal control decisions (policies) can come only from models which are not finite in size (again assuming there is some small region around the origin of state space in which linearized models are valid), alternately, that the simplest feedback control policies require no model at all (corresponding to open loop control)**. This is an easy

_____

*The actual arrangement of a set of models in order of their fidelity would be a most difficult task in itself owing to the fact that as the models become larger (i.e., as more modes of the system are included) more uncertainty is associated with the parameters.

**The existence of the hierarchy of models in Figure 1-1 has been argued. However, the actual construction of the "arbitrarily good" model mentioned above is impossible, for obvious reasons.

way to visualize the fact, well known to users of optimal control theory, that optimal solutions are only as good as the models upon which they are based. Thus, in solving optimal control problems there is always the (often unstated) economic constraint imposed on model size. In problems of analysis this constraint may be imposed by the capacity or cost of available computers. In problems of feedback control design this constraint may be imposed by the cost of synthesizing the control policy. The model size versus controlled system performance dilemma leads us to the characterization in section 2.0 of a problem we would like to solve (but cannot) called the "Minimal Controller Problem." Because the Minimal Controller Problem cannot be solved, it is typical to employ various schemes for "discrediting" the model (upon which control design is based) and these past approaches are discussed in section 3.0.

The approach taken in this research begins in section 4.0 and centers around the question, "Can an n_j vector, $e(t)$, (called a model error vector) be found such that when $e(t)$ is added to the state equations of any model $\mathcal{A}_j$ (see Figure 1-1), the output of the model $\mathcal{A}_j$ is caused to be identically equal to the output of a better model, $\mathcal{A}_1$?". Pursuit of this question (and the adding of precision to it) leads one to describe model error effects in four categories:

1. truncated modes

2. parameter errors

3. neglected external disturbances

4. neglected nonlinearities.

Characterizing an error vector e(t) which compensates for the above effects is not a straightforward task because one does not usually know, prior to design of the control policy, what disturbing effects (from external disturbances and truncated modes) are important to keep in the model. The control designer must typically rely upon performance evaluations in a trial and error fashion to determine what disturbance effects and modes of the dynamical system should be actively controlled, and therefore be included in the model upon which the control law is based. Also parameter uncertainties (whether constant are due to "in-flight" changes) and variations in the disturbance environment cannot usually be reliably predicted. It is for these several reasons that the appropriate model error vector, e(t), cannot be rationally specified prior to control system design. Therefore, the modeling problem is proposed as a two-phase task, as shown in Figure 1-2.

STEP I. Coarse model reduction ($\mathcal{A}_1 \rightarrow \mathcal{A}_3$ in Figure 1-2 describes a reduction from a conservative model, $\mathcal{A}_1$, to a reduced model $\mathcal{A}_3$ and is accomplished prior to control design).

STEP II. Vernier model adjustment ($\mathcal{A}_3 \rightarrow \mathcal{A}_2$ in Figure 1-2 describes a small increase in the order of the model by the augmentation of an "error system" to the model $\mathcal{A}_3$ and must be accomplished during the control design task. It will be shown that control designs based upon model $\mathcal{A}_2$ are less sensitive to modeling errors than designs based upon model $\mathcal{A}_3$.

This research is more concerned with a theory of modeling of linear systems than with any new theories of control, although from

6
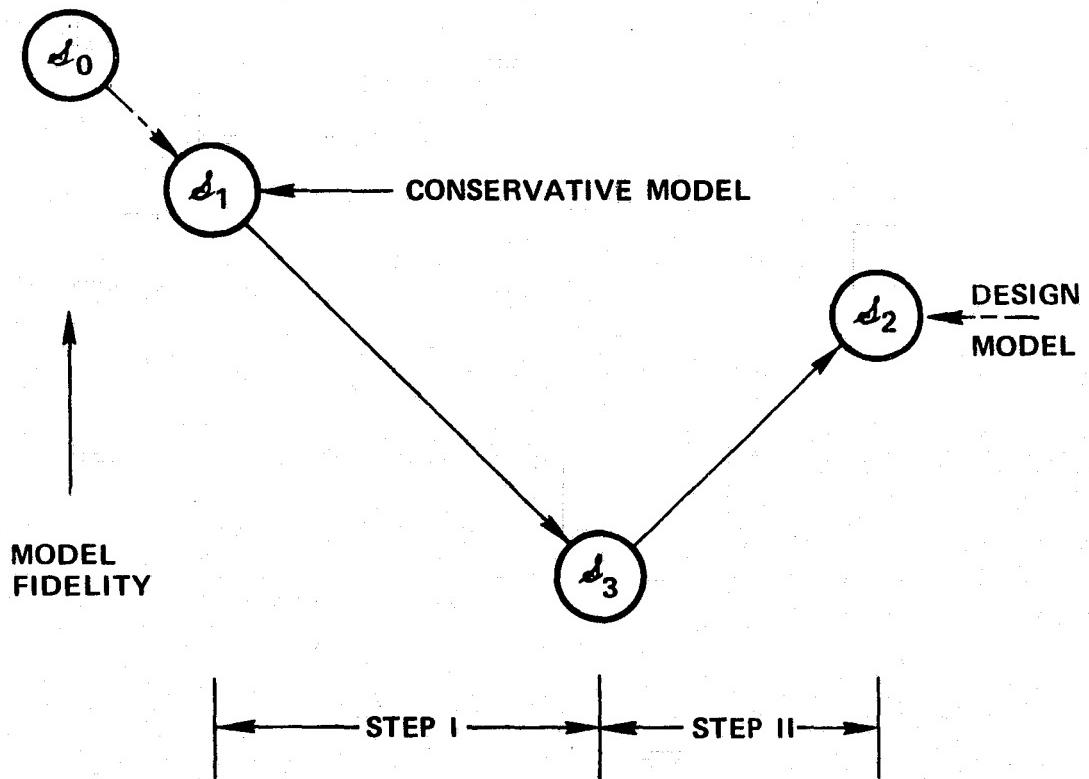
Figure 1-2. Design Model Development.

prior discussions we know that the two cannot be considered entirely separately. Therefore, section 4.0 will discuss the idea of control by model error estimation when the control problem of interest is the linear regulator problem of optimal control. In section 5.0 that particular choice for the model "error system" is made such that the model error vector considered evolves from a set of orthogonal functions which are considered to span the space of anticipated errors. A brief outline of the design procedure is given in section 5.0. This method of control design is applied to the Large Space Telescope (LST) which the National Aeronautics and Space Administration (NASA) plans to orbit in the 1980's.

The relationships of various aspects of this work with existing theories and practices are cited throughout the text, but a summarizing discussion of this sort is given in section 7.0, along with conclusions and suggestions for further research.

2.0   Some Effects of Model Errors in Control Problems

As modern control theory rapidly developed in the '60's the press of the control theorists to alleviate system performance limitations imposed by the control law (which he might have viewed as his responsibility) exposed, sometimes dramatically, the effects of errors in the system model (which he might have viewed as the dynamicist's responsibility). Such model error effects were not always discovered before "flight" of the system. Since failures of modern control implementations do not usually make the journals one can be mislead, on the basis of the literature, about the frequency of successes of modern control theory applications. Optimal control theory "squeezes" the models hard, and a theory that tells how to get more out of the system (model) puts greater pressure on the modeling process. Thus, the practicing engineer quickly learns that "optimal" solutions are only as good as the model upon which the solutions are based. Furthermore, since models for physical dynamical systems are always approximations, we must conclude that truly "optimal" decisions can only come from truly "best" models (since there always exists a better model, here "best" at least implies not finite dimensional). Therefore, practical questions of optimality carry an often unstated consideration of implementation constraints.[*] This leads us to the notion of the problem we would most like to solve (but cannot).

---

[*] In the mathematical sense the word "optimal" always means with respect to the given model and with respect to the chosen performance criteria. Poor system performance can also be attributed to a poor choice of performance criteria rather than inadequate modeling.

The Minimal Controller Problem:

Find a control policy of minimal complexity which allows the required performance to be guaranteed. That is, find the control policy $u^o = u^o(z^o, A^j, B^j, C^j, M^j)$ which although based upon real measurements z^o and the model, $\mathcal{A}_j$ (of the physical system, $\mathcal{A}_0$), satisfies the performance requirement

$$v^j \triangleq \int_0^\infty (||y^o||_Q + ||u^o(z^o, \mathcal{A}_j)||_R) dt \leq v^o$$

with minimum n_j, where V^o is a specified number and where

$$\mathcal{A}_0 \triangleq \text{physical system} \begin{cases} u^o: & \text{actual inputs } \in R^m \\ y^o: & \text{actual outputs } \in R^k \\ z^o: & \text{actual measurements } \in R^\ell \end{cases}$$

$$\mathcal{A}_j \begin{cases} \dot{x}^j = A^j x^j + B^j u^o \\ y^j = C^j x^j \\ z^j = M^j x^j \end{cases}$$

x^j: state $\in R^{n_j}$

u^o: control $\in R^m$

y^j: output $\in R^k$

z^j: measurement $\in R^\ell$

Note that z^o is a real measurement and cannot be predicted with certainty. Now therefore the Minimal Controller Problem is not a mathematical problem (due to the requirement of certainty, $v^j \leq v^o)^*$ and cannot be solved mathematically, and for that matter we have no procedure to solve it even empirically. However one might attempt to solve the minimal controller problem in the following way. Presume that the output variables y^o which we desire to control have been identified and that the "controller," $u^o(\mathcal{A}_j, z^o)$, is composed of the

---
*The problem cannot be solved even in the statistical sense; see section 3.0.

control law $u^o = Gx^j$ and the state estimator

$$\dot{\hat{x}}^j = [A^j - \hat{K} M^j]\hat{x}^j + B^j u^o + \hat{K} z^o.^*$$

a. Choose the decision variables, u_i^o (with due regard to the fact that the model $\mathscr{A}_j$ presumes that $u_i^o(t)$ can be instantaneously changed. In practice any delays in changing u must be small with respect to the time constants associated with both the open loop and closed loop system of model $\mathscr{A}_j$).

b. Choose a model order, n_j (for the sake of this discussion begin with $n_j = 1$).

c. Choose model parameters (A^j, B^j, C^j, M^j) (Note that this involves, in (A^j, M^j), decisions of what measurements to make and in (A^j, B^j) decisions of where to apply the controls).

d. Choose a controller (that is, choose a functional relationship between the measurements, $z^o(t)$, and the control issued, $u^o(t)$).

e. Evaluate V^j and test the inequality $V^j \leq V^o$. (If yes stop. If no, return to c. or d. After exhausting "all" possibilities in c. and d., return to b. increasing n_j by 1. Continue until e. is satisfied.)

The statement of the minimal controller problem serves only to remind us of desired objectives and is not in itself constructive. Indeed, it is because we cannot solve the Minimal Controller Problem that we are forced to pay close attention to model errors.

Before considering the effects of model errors we should define the terms. Errors in linear models may result from;

---

*The details of estimator (full order and reduced) design will be discussed in sections 3.0 and 4.0.

11

1. Truncated modes (those errors introduced by model reduction, recognizing that _any_ mathematical model may be considered a reduction from a better one; thus this effect is _always_ present in models of physical dynamical systems.)

2. External disturbances (those unmodeled effects which are independent of the state variables used to describe the model).

3. Parameter errors (even if the structure (order) of the model is appropriate, there may exist better values for the parameters).

4. Effects of linearization (neglected nonlinearities).

Now, we wish to know now what price we pay for the presence of such modeling errors.

2.1 Some Effects of Model Errors in Deterministic Characterizations of Dynamical Systems.

Suppose we wish to minimize the functional

$$J(u^o) = \frac{1}{2} \int_0^T \left(y^{j^T} Q y^j + u^{o^T} R u^o \right) dt + \frac{1}{2} y^j(T)^T H y^j(T) \qquad (2.1)$$

subject to the time-invariant linear system model, $\mathscr{A}_j$

$$\mathscr{A}_j \begin{cases} \dot{x}^j = A^j x^j + B^j u^o & x^j \in R^{n_j} \\ \begin{pmatrix} y^j \\ z^j \end{pmatrix} = \begin{bmatrix} c^j \\ M^j \end{bmatrix} x^j & y^j \in R^k \qquad (2.2) \\ x(t_o) = x_o \end{cases}$$

Furthermore, let the pair (A^j, B^j) be controllable[*] and (A^j, C^j) be observable[**] in the sense of Kalman (also, let Q, R, be positive definite matrices). Then, according to this model, the closed loop optimal system obeys

$$\dot{x}^j = \left[A^j - B^j R^{-1} B^{j^T} K \right] x^j = \bar{A} \, x^j \tag{2.3}$$

where K is the symmetric, positive definite solution of the matrix Riccati equation

$$\dot{K} = - K A^j - A^j K + K B^j R^{-1} B^{j^T} K - C^{j^T} Q C^j \tag{2.4}$$
$$K(T) = H$$

Now $\bar{A}$ has an arbitrary spectrum. That is, there exist real values for Q and R such that the eigenvalues of $\bar{A}$ may be arbitrarily specified,[13]. Then, on the basis of the analysis of this system model, arbitrarily good performance can be predicted. That is, there exist Q, R, such that

$$J(u^o) \leq \varepsilon \qquad \forall \varepsilon \geq 0 \tag{2.5}$$

To synthesize (operate the <u>real system</u>) with this "optimal" control policy, $u^o = - R^{-1} B^{j^T} K x^j$, a "separation principle" is usually invoked, as follows. The control law is determined (above) as if the state x^j

_____

[*] The pair (A^j, B^j) is controllable if rank $[B^j, A^j B^j \ldots A^{j^{n_j-1}} B^j_1] = n_j$.

[**] The pair (A^j, C^j) is observable if rank $[C^{j^T}, A^{j^T} C^{j^T}, \ldots, A^{j^{T^{n_j-1}}} C^{j^T}] = n_j$.

In this report the word "controllable" will always mean with respect to the state vector (that is (A^j, B^j), satisfy above condition.) Output controllability (rank $[C^j B^j, C^j A^j B^j, \ldots, C^j A^{j^{n_j-1}}]$ $[C^j B^j, C^j A^j B^j, \ldots, C^j A^{j^{n_j-1}} B^j] = n_j$) is assumed throughout.

were available for measurement. A state estimator is then constructed

to asymptotically reconstruct x^j. The estimate of x^j, labeled $\hat{x}^j$, is

then used to synthesize the feedback control solution,

$u^o = - R^{-1}B^{j^T}K\hat{x}^j$. The justification offered for this approximate

implementation of the optimal solution is that the state estimator (or

state "observer" as the literature on deterministic problems prefers)

can, under certain conditions, (observability of the pair (A^j,M^j))

be made arbitrarily "fast" so that errors between the state, x^j, and

the estimate of the state, $\hat{x}^j$, are arbitrarily small after an arbi-

trarily short period of time. Thus, the synthesized "controller" is

as shown in Figure 2-1, where it is noted that the estimator has an

arbitrary spectrum (eigenvalues of $(A^j - \hat{K} M^j)$ may be arbitrarily

specified) by virtue of observability of A^j, M^j.* The full n^{th}-order

form of the state estimator is shown in Figure 2-1 instead of the

reduced, $(n_j - \ell)^{th}$-order form, which might be implemented in practice,

to illustrate more clearly the effects of model errors. If the control

gains G are taken as those optimal for the model $\mathscr{A}_j$ then

$G = - R^{-1}B^{j^T} K$, see (2.1) - (2.4). Now the problem that sometimes

occurs with the control scheme of Figure 2-1 is that in spite of the

fact that arbitrarily good performance can be predicted on the basis

of the model $\mathscr{A}_j$, the physical system $\mathscr{A}_o$ can deliver arbitrarily bad

performance (i.e., y^o diverges from desired value $y^o = 0$). The gen-

eral situation is that while "arbitrarily good" performance is not a

---

*See Wonham[58] for one proof of this fundamentally important result of
linear system theory which guarantees the existence of a real matrix
M which will give the matrix (L+MN) an arbitrarily specified set of
eigenvalues if and only if the pair (L,N) is observable.

PHYSICAL SYSTEM
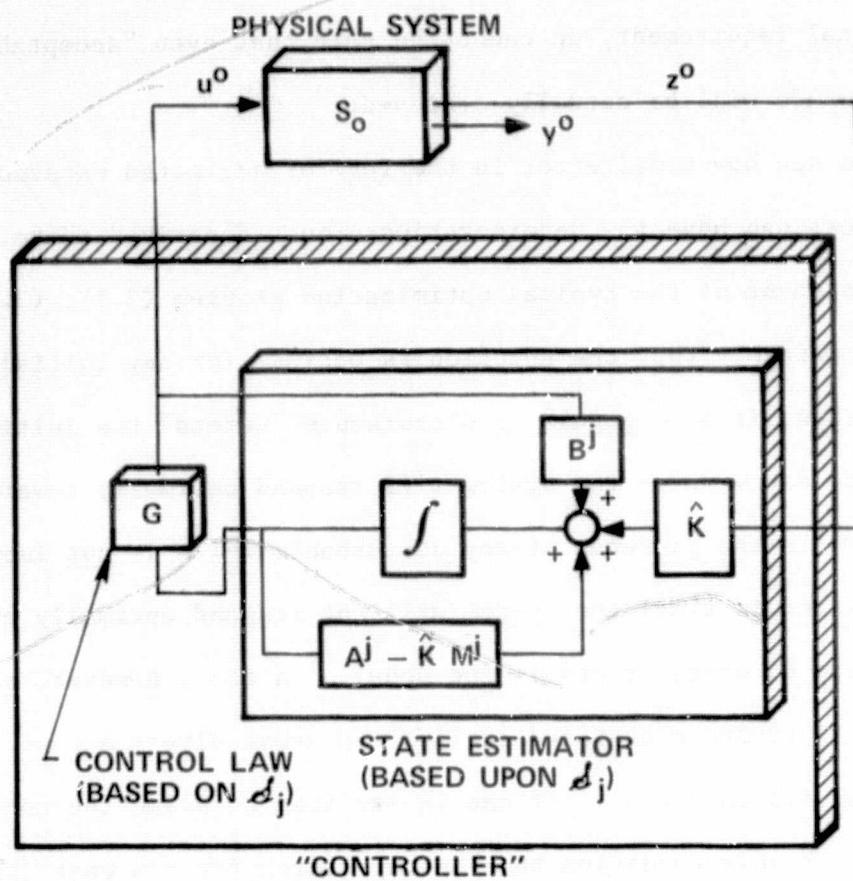
$$\mathscr{A}_j \begin{cases} \dot{x}^j \simeq A^j x^j + B^j u^o \\ \begin{pmatrix} y^j \\ z^j \end{pmatrix} = \begin{bmatrix} C^j \\ M^j \end{bmatrix} x^j \end{cases}$$

Model $\mathscr{A}_j$ is characterized by the parameters

$$(A^j, B^j, C^j M^j) : \text{MODEL PROBLEM}$$

State estimator is characterized by $(\hat{K})$: ESTIMATOR PROBLEM

Control Law is characterized by (G) : CONTROL PROBLEM

Figure 2-1. The Controller for the Linear Regulator Problem.

15

practical requirement, we cannot be sure that even "acceptable" performance will be actually achieved.

To see how model error in the form of neglected external disturbances can have the deteriorating effect described above, consider the statement of the typical optimization problem (2.1), (2.2). We see immediately that the solution is optimal for any <u>initial condition</u>, x_o^j . Thus, if some <u>impulsive</u> disturbance "resets" the initial conditions to a new value the system will respond optimally toward $y^j = 0$. However, in the presence of any disturbance which is <u>not</u> impulsive (correlated in time) the system will not respond optimally even if there are no other errors in the model. In fact, however, each of the model errors classified earlier can cause divergence of the physical system output. If one is inclined to blame the control theory for this condition then he may search for new ones (such as the adaptive control theories). If one is inclined to blame the modeling process for this condition, then he may search for better models. This latter attitude is presented in this research.

A few words are needed here to justify our attention wholely to state space models and further, to state space models which are not minimal realizations.[*] Our rationale rests solely on the argument that physical dynamical systems are not completely state controllable. To be more precise we must argue that as we progress up the hierarchy of (mathematical) models in Figure 1-1 (which, recall, is arranged in order of fidelity of the models), the expectation of controllability

---

[*] Minimal realizations are state observable, controllable models, see Reference 40, p. 94.

16

of the pair (A^j, B^j) decreases. This notion is at odds with those offered by Ho and Lee,[14] Kreindler and Sarachik,[15] and Schultz and Melsa, Reference 16, p. 29, who suggest that uncontrollable models of physical systems are "inadvertent." To support our claim to the contrary consider the

PROPOSITION: For every completely controllable linearized model of a physical dynamical system, there always exists an uncontrollable linear model of the physical system which more accurately portrays the outputs of the physical system.

For supporting arguments we cite two separate circumstances, the first based upon the presence of external disturbances, the second on internal disturbances. First consider that for any model of the system there can always be identified some disturbance which is "external" in the sense that it does not depend upon the state variables previously in the model* (i.e., the neglected gravity gradient torques on a spacecraft, or thermal gradients on a ground based antenna etc.). Now adjoin a differential equation which that disturbance obeys (finding that equation is quite another task) to the system description. The resulting composite system then provides a more faithful characterization of the output, $y(t)$, than did the model which neglected the disturbance. Moreover, those variables associated with the model of the disturbance are uncontrollable. The second circumstance of

_____

*Thus, whether a particular effect qualifies as an external disturbance depends upon the particular model, $\mathcal{A}_j$, at hand. Note that such disturbing effects fall in either the category of external disturbances or truncated modes, as listed on page 12.

"internal disturbances" (truncated modes) is discussed in detail by Likins and Ohkami.[66] In this report it is shown that some modes may be found in a linear model of a physical dynamical structure, such as a spacecraft, which are uncontrollable.

Now the danger in using minimal realizations to describe physical processes is that while under controlled conditions (i.e., laboratory prototypes) a controllable model may adequately describe the response of the output, under "flight" conditions the natural environmental disturbances (and truncated modes) may excite uncontrollable modes (which are present in the physical system[*]). In this circumstance the controllable models are loathe to explain the discrepancies observed in the measurements and hence even adaptive techniques, designed to update parameters within a controllable model, may fail to yield stabilizing feedback. Thus, the choice of words "minimal realization" to describe controllable, observable models of minimal order is per-haps unfortunate since they should <u>not</u> be taken to mean that if a model is not controllable there is another, equally credible, model that is.

It is not surprising then, that Astrom[1] has found that increasing the order of a model can compensate for parameter errors (without the use of adaptive techniques). The converse is not true. That is, changes in parameters may not compensate for the effects of truncated

---

[*] The phrase "physical system" necessarily includes effects which qualify as "external disturbances" with respect to a simpler (mathematical) model. See footnote on page 17.

modes. This point is discussed in section 3.0. Concerning uncontrollable modes then we conclude that they may be (automatically) truncated from the model only to the extent that they are also unobservable.

Finally, we note that classical techniques are restricted to controllable (and observable) models and cannot be used to directly compensate the effects of uncontrollable modes and external disturbances. (This is one reason why classical design procedures are often characterized in the oversimplified terms "design for stability, then check for performance" and optimal control design procedures in the terms "design for performance, then check for stability"). Compensation of external disturbances may be (indirectly) accomplished via classical techniques if the right structure for the compensation is proposed a priori. For example, if a bias disturbance is noticed (in the "check for performance" phase of design), by the appearance of a bias error in a pointing control problem, then an integrator can be added in the compensation to correct for this disturbance. We will, instead, use a procedure which allows consideration of both uncontrollable modes and the treatment of disturbances directly.

One criticism which has been leveled at control policies which utilize state estimators (see Figure 2-1) is "Why estimate all of the state variables when only some might accomplish a satisfactory control policy?" Our response to this question is that if we consider ourselves to be responsible for both the model and the control law design, then, out of concern for the Minimal Controller Problem (presented in section 2.0), we take the attitude that if a particular set of state

variables does not need to be estimated to accomplish the control

task then those state (modes) do not belong in the model to be used

for controller design.

The overall control design task is often considered, from Figure

2-1, as three separate problems:

1. Selection of (A^j, B^j, C^j, M^j) (including dimensions as well

as parameter values) as the modeling problem,

2. Selection of $\hat{K}$ as the estimator problem,

3. Selection of G as the control problem.

(Alternately, items 2 and 3 taken together may be considered as the

controller problem).

In this research we will consider how these three problems might

be related through an approximation of the model errors which are

inherent in any finite dimensional model of a physical dynamical

system.

It is noted that the matrix $\hat{K}$ in Figure 2-1 can be chosen accord-

ing to a number of different criteria. If the problem description is

deterministic, as above, then $\hat{K}$ is chosen for desirable stability

properties of the estimator. If the problem description is stochastic

the $\hat{K}$ can be fixed in terms of statistical properties of the random

variables of the problem. Thus as a special case Figure 2-1 also

describes the Kalman filter which is discussed in the next section.

2.2 Some Effects of Model Errors in Stochastic Characterizations of
     Dynamical Systems

In the previous section the point was made that model errors in

deterministic models can cause state estimators to diverge, resulting

in instability of the closed loop system. One approach often used to provide a "dynamical cover" for the model errors is to add a noise "disturbance" term w(t) to the model and write

$$\dot{x} = A x + B u + w(t)$$

so that the ensuing mathematics will not lead us to believe that the model (of the undisturbed system) is "arbitrarily good." In other words some noise (real or imagined) can be added to "discredit" the previous model. Now the point to be made in this section is that if the actual disturbance[*] acting on the system is a <u>correlated</u> random process and we have modeled the disturbance w(t) as a <u>white</u> noise process then the resulting state estimator may still diverge (further details will follow in section 3.0). Moreover, there may not exist <u>any</u> set of parameters within the fixed structure of the model or any set of parameters for the covariance matrices of the random variables (only first and second order statistics are needed for the optimal linear state estimator) which will yield a stable estimator.

To illustrate this problem let us refer to Figure 2-2 which shows the solution of the linear quadratic Gaussian problem with the resulting Kalman filter identified. From the Figure it is seen that the model $\mathcal{A}_j$ of the physical system $\mathcal{A}_o$ is shown with the same notation (the superscript j is chopped in Figure 2-2 for clarity) as in the previous section with the exception that the initial condition x_o, the plant disturbance, w(t), and the measurement noise v(t) are all

_____

[*]It will be shown in section 4.0 that there exists a "model error vector," e(t), which when added to the model (i.e., $\dot{x} = A x + B u + e(t)$) can correct for model errors due to a) truncated modes (unmodeled dynamics), b) external disturbances, c) parameter errors, and d) weak nonlinearities. In this section e(t) is considered to be white noise.

PHYSICAL SYSTEM

CONTROL LAW
(BASED UPON $\mathscr{S}_j$)

STATE ESTIMATOR (KALMAN FILTER)
(BASED UPON MODEL $\mathscr{S}_j$)
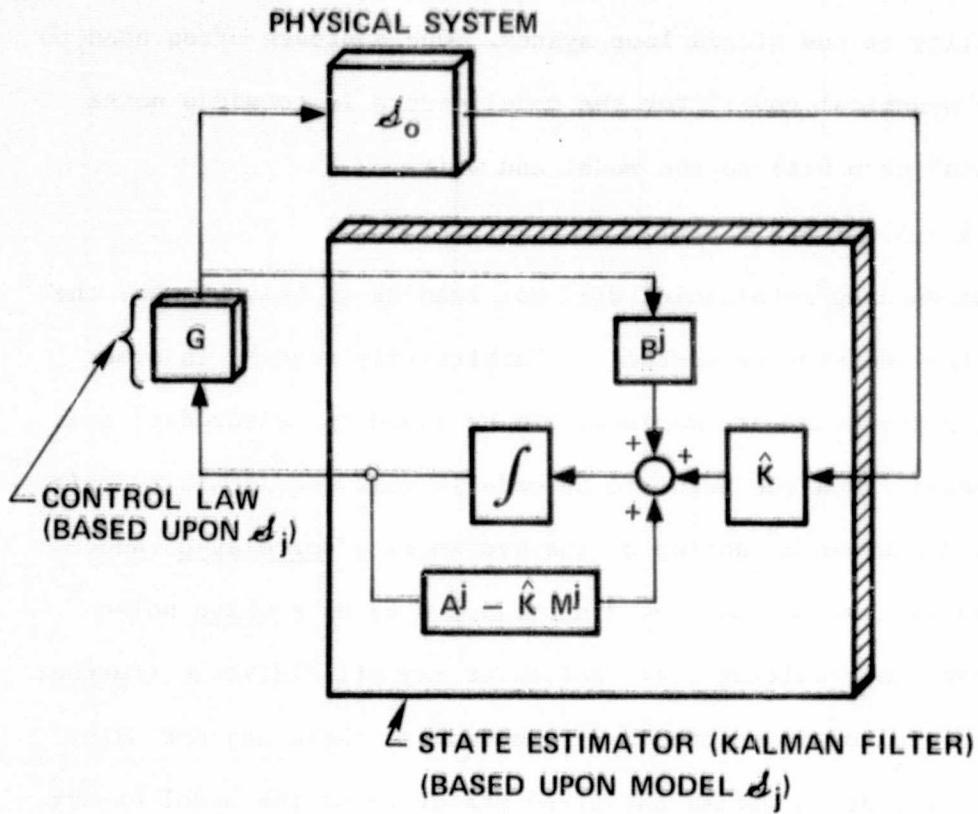
$$\mathscr{S}_j \begin{cases} \dot{x}^j = A^j\,x + B^j\,u + w(t) & w(t) \sim N(0,\, \hat{Q}(t)) \\ z = M\,x + v & v(t) \sim N(0,\, \hat{R}(t)) \end{cases}$$

Control $u = -R^{-1}B^{j^T} K\,\hat{x} = G\,\hat{x}$ \qquad $x_o \sim N(0,\, P_o)$

where K obtains from (2.4), and \qquad $E[w(t)w^T(\tau)] = \hat{Q}(t)\,\delta(t-\tau)$

$$\dot{\hat{x}} = A^j\,x + B^j\,u^o + \hat{K}(z^o - M^j\,\hat{x}) \qquad E[v(t)\,v(\tau)] = \hat{R}(t)\,\delta(t-\tau)$$

where $\hat{K} = \Sigma\,M^{j^T}\hat{R}^{-1}$ \qquad $E[w(t)\,x_o^{\,T}] = 0$

$$\dot{\Sigma} = A^j\Sigma + \Sigma\,A^{j^T} - \Sigma M^{j^T}\hat{R}^{-1}M^j\Sigma + \hat{Q} \qquad E[w(t)\,v^T(t)] = 0$$

$$\Sigma(t_o) = \Sigma_o \qquad\qquad E[v(t)\,x_o^{\,T}] = 0$$

$$E[x(t)x^T(t)] \;\underline{\Delta}\; \Sigma(t)$$

$$E[x_o\,x_o^{\,T}] = \Sigma_o$$

Figure 2-2. Linear Quadratic Gaussian Problem.

Gaussian random vectors with zero means and covariances P_o, $\hat{Q}(t)$, and $\hat{R}(t)$, respectively. Furthermore, $v(t)$ and $w(t)$ are uncorrelated (white) processes. Now even though we assume the model $\mathcal{A}_j$ is completely controllable and observable in the stochastic sense, the Kalman filter (state estimator) may still diverge for the same reasons as in the deterministic case: model infidelity. Consider the particular example of very small, or zero, plant noise $\hat{Q} \approx 0$. Now, the estimator gain, $\hat{K}$, which was non-uniquely chosen for stability of the estimator in the deterministic problem of the previous section, (compare Figures 2-1 and 2-2) is now uniquely determined in terms of the statistical parameters of the random vectors x_o, $w(t)$, and $v(t)$ (as shown in the equations for $\hat{K}$ in Figure 2-2), and $\hat{K}$ is called, in this case, the Kalman gain. Since $\hat{Q} \approx 0$, $\hat{K}$ has a zero steady state solution, due to the fact that $\Sigma = 0$ is a steady state solution of the covariance equation. The result is filter divergence since for $\hat{K} = 0$ the filter is no longer causally related to the measurements. Thus, the filter "thinks" it has "learned" the model and disconnects itself from the noisy measurements (by making $\hat{K} \to 0$) and relies only on predictions from the (always erroneous) model $\mathcal{A}_j$. Divergence occurs due to the fact that undue confidence is placed on the model (here the word "model" includes the statistical parameters $\hat{Q}$, $\hat{R}$).

This example illustrates how the gain $\hat{K}$ serves as an "arbitrator" to weight between the measurements (which evolve from the _real_ system $\mathcal{A}_o$) and the predictions (which evolve from the selected _model_ $\mathcal{A}_j$ of the real system $\mathcal{A}_o$). The estimator "decides" which to believe (weight more heavily) based upon model parameters presumed ($\hat{Q}(t)$, $\hat{R}(t)$,

Σ_o, A,B,M). Thus, when the model is inadequate, the "decisions" made by the filter can be bad, leading to system instability. This example also has led many researchers to seek a "fix" for this divergence problem by increasing $\hat{Q}$ in some way,[18-20]* thereby artificially, or otherwise, decreasing the confidence in the model by causing $\hat{K}$ to be increased. Detailed conditions of Kalman filter divergence may be found in References 21 through 24.

Conversely, the theorem of Price[21] gives the conditions for uniform asymptotic stability in the large. Under stated conditions the theorem assures that the norm of the covariance matrix has an exponential upper bound

$$||E[(x^{j'} - \hat{x}^j)(x^{j'} - \hat{x}^j)^T]|| < \sigma_1 e^{-\sigma_2 t}$$

where $x^{j'}$ is the state of a better model of the same dimension (i.e., more correct parameters) and $\hat{x}^{j'}$ is the state of the model $\mathscr{A}_j$ upon which the filter design is based. The unsettling feature of this result is that it is still possible for the filter to satisfy these conditions (and thus be uniformly asymptotically stable in the large) and yet, <u>within time constants of the system</u>, yield arbitrarily bad estimates. This can occur because no upper bound is established for σ_1 and σ_2.

In a similar manner, persistent unmodeled "disturbances" (which may result from truncated modes as well as external disturbances) do the same "damage" to a deterministic (Luenberger) observer as unmodeled

---

*Fitzgerald,[22] shows that an increase in $\hat{Q}$ can also make the estimates worse.

correlated disturbances do to the Kalman filter. We now refer to some specific alternatives and past attempts to compensate for the presence of model errors in the control design problem.

3.0 Past Approaches to Compensation for Model Errors

What follows in this section is a brief summary of various methods
of control design in the presence of modeling errors of the type dis-
cussed in section 2.0. A point of view one can adopt to help unify
these approaches is to consider that the basic objective common to all
methods to be discussed is to somehow "discredit" the mathematical
model. After all, discrediting an erroneous model is a first step to
better control decisions.

Basically, the problem is as follows: We have committed the
physical system to a linear model, $\mathscr{A}_j$

$$
\mathscr{A}_j \quad
\begin{cases}
\dot{x}^j = A^j x^j + B^j u^o \\
y^j = c^j x^j \\
z^j = M^j x^j
\end{cases}
\tag{3.1}
$$

where z^j represents measurements (as predicted by the model, $\mathscr{A}_j$)
where u^o represents the "decision," or "control," variables selected
for manipulation, and y^j represents (as well as this model is able to
predict) the output variables, the object of our control. We wish the
variables $y^j(t)$ to behave in a certain way in accordance with some
mathematical criteria we have specified to reflect the desired per-
formance of the physical system. Now, the selection of $u^o(t)$ is to be
based upon analysis of the model $\mathscr{A}_j$ (to avoid trial and error methods
with the physical system). On the basis of this "analysis" (the pre-
diction of physical system performance on the basis of the model, $\mathscr{A}_j$)
we can provide $u^o(t)$ in an open loop fashion,

PRECEDING PAGE BLANK NOT FI.

$$u^o(\mathcal{A}_j, x_o^j, t) \longrightarrow \boxed{\mathcal{A}_o} \longrightarrow y^o(t)$$
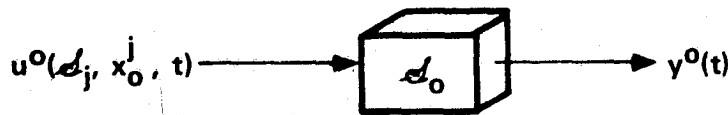
closed loop fashion (if some measurements $z^o(t)$ are available),
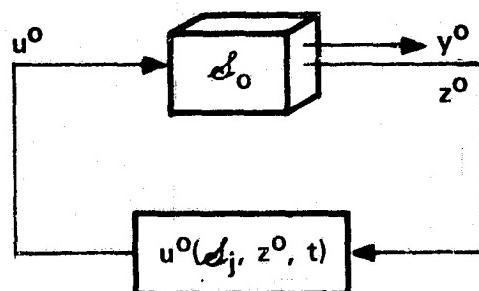
and modifications of the closed loop fashions which include the

capability to change parameters within the fixed structure of $\mathcal{A}_j$.
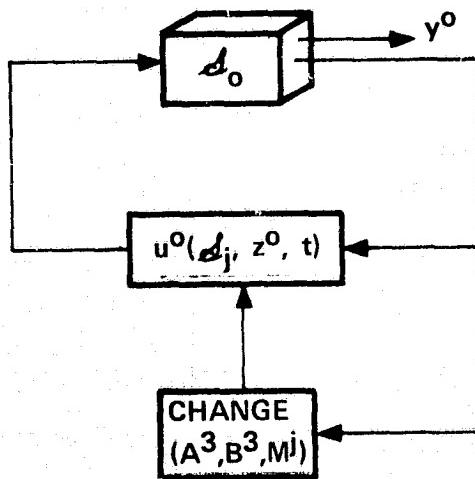
This latter scheme falls under the category of "adaptive" control.

Now let us begin our brief survey.

3.1 Feedback Control

If the model $\mathcal{A}_j$ was a perfectly accurate representation of the physical system then the physical output $y^o(t)$ would behave in exact agreement with the output $y^j(t)$ as predicted on the basis of analysis with the model $\mathcal{A}_j$. In this event, one can issue the controls $u^o(t)$ to the physical system in an open loop manner with confidence that the physical output $y^o(t)$ will behave exactly as predicted. However, the effects of nonlinearities, parameter errors, and external disturbances, and modes of the dynamical system which have not been modeled (the collection of these effects has been labeled in section 2.0 as "modeling errors") cause physical systems to deviate from their modeled behavior. It is primarily for this reason that feedback control has been so useful. Even though the selection of gains or dynamical elements in the feedback path is still based upon analysis with model $\mathcal{A}_j$, the current measurements from the physical system serve to automatically "update" the initial conditions presumed for the model. The open loop command, on the other hand, is based upon prediction of system performance for all future time from the fixed initial conditions presumed for the system. Thus, feedback has the effect of shortening the time over which the model $\mathcal{A}_j$ needs to provide accurate predictions.

3.2 Conservative Designs

To partially compensate for uncertainty in the model a "safety margin" can be incorporated into the performance specifications. Uncertainty in the disturbance environment may lead one to use "worst case" disturbances for control system evaluation purposes. Similarly,

"minimax" designs have been used,[25] to obtain the minimum of a performance criterion subject to those particular parameter changes which tend to maximize the criteria.

3.3 Parameter Sensitivity Approaches

The attitude taken in sensitivity approaches,[26-34] is that if the values of the parameters of the model are uncertain then one should add sensitivity measures to the performance criterion to be minimized. Sensitivity of the original performance criterion or sensitivity of the state trajectories may be considered. In the latter case the new optimization problem is:

Minimize
$$J(u) = \int_0^T (x^T Q x + s^T S s + u^T R u) dt$$

$$\dot{x} = A(p)x + B(p)u \qquad (3.2a)$$

subject to
$$\dot{s}^1 = \frac{\partial A(p)}{\partial p_1} x + A(p) s^1 + \frac{\partial B(p)}{\partial p_1} u + B(p) \frac{\partial u}{\partial p_1}$$

$$\vdots \qquad (3.2b)$$

$$\dot{s}^r = \frac{\partial A(p)}{\partial p_r} x + A(p) s^r + \frac{\partial B(p)}{\partial p_r} u + B(p) \frac{\partial u}{\partial p_r}$$

where $s^i \triangleq \frac{\partial x}{\partial p_i}$, $\frac{\partial s}{\partial p} = 0$ (assume)

$$u = F_1 x + F_2 s$$

and Q , S , R are positive definite matrices.

Newmann,[32] shows that this problem generally has no solution and Quardabassi,[33] proves why; the system is uncontrollable for a sufficiently large number of parameters (r). Some results can be obtained if r is small and the requirements of positive definiteness of S are relaxed. Sakharov[34] suggests that, in general, it is necessary

30

to use first and second order sensitivity models. However, if the

trajectory x(t) is itself erroneous because of model errors then

differentiation of x(t) (with respect to p) provides variables,

$s = \frac{\partial x}{\partial p}$, which seem to provide even less reliable information about the

system. Differentiating x twice (to obtain second order sensitivity)

seems even more hazardous. Also note that the state vector is <u>doubled</u>

in size for <u>each</u> parameter considered uncertain in the original model.

3.4 Singular Perturbation Approaches

The sensitivity methods above increase the order of the system

under study. That subsystem (3.2b) which was augmented to the model

(increasing the order) was based solely upon information contained in

the smaller model, $\mathcal{A}_j$. An alternate approach is to consider that

every model of a physical system can be considered a truncation from a

better model. In this view it is possible to increase the order of

$\mathcal{A}_j$ by adding to $\mathcal{A}_j$ those modes which were truncated to obtain $\mathcal{A}_j$

originally. Now the optimal control solution for this higher order

model can be approximated by the application of singular perturbation

methods.[35] The method applied to the linear regulator problem proceeds

as follows, see References 7 through 9. Presume that the equations

can be put in the form

$$\left\{ \begin{array}{c} \dot{x} \\ \varepsilon \dot{x}^t \end{array} \right\} = \begin{bmatrix} A_{11}(\varepsilon) & A_{12}(\varepsilon) \\ A_{21}(\varepsilon) & A_{22}(\varepsilon) \end{bmatrix} \left\{ \begin{array}{c} x \\ x^t \end{array} \right\} + \begin{bmatrix} B_1(\varepsilon) \\ B_2(\varepsilon) \end{bmatrix} u \qquad (3.3)$$

where ε is a small parameter and where x is an n-vector of variables.

Now x^t is an N-vector (N might be large) of variables we would like to

truncate from our model, but because the truncated model

31

$$\mathcal{A}_j \left\{ \dot{x} = \tilde{A}\, x + \tilde{B}\, u, \text{ let} \begin{cases} \tilde{A} \triangleq A_{11}(0) - A_{12}(0)A_{22}^{-1}(0)A_{21}(0) \\[2mm] \tilde{B} \triangleq B_1(0) - A_{12}(0)A_{22}^{-1}(0)\, B_2(0) \end{cases} \right\} \qquad (3.4)$$

might yield a controlled system which is sensitive to the neglected effects of x^t, we instead wish to find an <u>approximate</u> solution of the control problem using the higher order model (3.3). The exact solution for

$$\min_u \; J(u) = \int_0^T (x^T Q x + u^T R u)\, dt + x(T) H\, x(T) \qquad (3.5)$$

subject to (3.3) and $x(0) = x_o$; $x^T \triangleq (x, x^t)$; Q,F, positive semi-definite and R positive definite, is

$$u = -\, R^{-1} B^T K\, x \qquad (3.6)$$

where K obtains from

$$\dot{K} = -\, K A - A^T K + K B R^{-1} B^T K - Q \qquad K(T) = H \qquad (3.7)$$

and

$$A = \begin{bmatrix} A_{11}(\varepsilon) & A_{12}(\varepsilon) \\[4mm] \dfrac{A_{21}(\varepsilon)}{\varepsilon} & \dfrac{A_{22}(\varepsilon)}{\varepsilon} \end{bmatrix} \qquad\qquad B = \begin{bmatrix} B_1(\varepsilon) \\[4mm] \dfrac{B_2(\varepsilon)}{\varepsilon} \end{bmatrix} \qquad (3.8)$$

$$Q = \begin{bmatrix} Q_1 & 0 \\[3mm] 0 & 0 \end{bmatrix} \qquad\qquad H = \begin{bmatrix} H_1 & 0 \\[3mm] 0 & 0 \end{bmatrix}$$

Thus a Riccati equation for the $(n+N)^{th}$ order system must be solved. Now, in Reference 9 the approximation suggested is that the control gain matrix $G = -\, R^{-1} B^T K$ be expanded in a MacLaurin series in ε and the first two terms be kept.

32

$$G(\varepsilon) \cong G(0) + \left[\frac{\partial G}{\partial \varepsilon}(0)\right]\varepsilon$$

This result leads to a requirement to solve $(N(N+1)/2 + n(n+1)/2)$ linear scalar equations for G (and inverting an NxN matrix) instead of the $(n+N)(n+N-1)/2$ nonlinear scalar equations required in the exact high order solution.

The method proposed in Reference 8 provides an alternate approximation by solving a boundary layer problem using two time scales, t and $\tau = \frac{t-T}{\varepsilon}$. The Riccati matrix K is then approximated by "slow" and "fast" solutions

$$K(t,\varepsilon) = \widetilde{K}(t,\varepsilon) + \widetilde{\widetilde{K}}(\tau,\varepsilon)$$

where $\widetilde{K}(t,\varepsilon)$ is a power series in ε and is a function of the slow variable, t, and where $\widetilde{\widetilde{K}}(\tau,\varepsilon)$ is a power series in ε and is a function of the fast variable, τ . The solution is then obtained by matched asymptotic expansions.[35] The number of scalar equations to be solved is the same as in the above case.

This solution of the problem handles higher order effects (from modes that might be truncated in a low order solution) very well. It is still susceptible to external disturbances and parameter errors. The implementation of the solution (to form the controller) is of the higher order.

3.5 Parameter Estimation Techniques

Suppose we have determined that the zero state response of a physical system can be reliably modeled by the completely controllable, observable linear model $\mathcal{A}_j$, of order n_j. Due to the fact that parameter changes, which are expected in the system from the aging of

components, the changing of the environment, etc., cannot be reliably predicted, we may wish to provide a means to update the parameters presumed within the fixed structure of the model $\mathcal{A}_j$. There are on-line and off-line techniques to accomplish this, Reference 1, and we mention here only one basic scheme.

If we label the uncertain parameters, $p_1, p_2, \ldots, p_r$, then state estimation of the augmented system

$$
\begin{cases}
\dot{x}^j = A^j(p)x^j + B^j(p)u^o \\
\dot{p} = 0 \\
z^i = M^j x^j
\end{cases}
\tag{3.9}
$$

in response to real measurements $z^o(t)$ may provide better values for p. This problem is difficult because it is now a nonlinear problem. The method also cannot account for other types of model errors such as external disturbances and truncated modes. Astrom[1] has shown that increasing the order of the (linear) model can compensate for parameter errors, without adaptive and nonlinear methods. Another way to increase the system order is discussed next.

3.6 External Disturbance Point of View

Instead of concentrating on errors in the parameters of model $\mathcal{A}_j$, suppose we consider the effect of an external disturbance $w(t)$ on $\mathcal{A}_j$

$$
x^i = A^j x^j + B^j u^o + w(t)
\tag{3.10}
$$

3.6.1 Deterministic Disturbance Characterizations

If we can construct a differential equation which $w(t)$ is known to obey

$$w = Pz$$

$$\dot{z} = Dz \tag{3.11}$$

then we can augment this disturbance system to $\mathcal{A}_j$ and obtain the new system

$$\begin{pmatrix} \dot{x}^j \\ \dot{z} \end{pmatrix} = \begin{bmatrix} A & P \\ 0 & D \end{bmatrix} \begin{pmatrix} x^j \\ z \end{pmatrix} + \begin{bmatrix} B^j \\ 0 \end{bmatrix} u^o \tag{3.12}$$

$$z^j = M^j x^j$$

which was studied by Johnson[2] and Johnson and Skelton,[3] yielding an approach which came to be known as "Disturbance accommodating Controllers," Johnson.[4] Davison,[6] Bhattacharyya and Pearson[5] also studied this problem. In fact, it was the writing of a differential equation which an input (command or disturbance) obeys that led to the solution of the servomechanism problem in Bhattacharyya and Pearson.[36] The "tracking" of a time varying command or the "rejection" of a time varying disturbance both reduce to a linear regulator problem once the differential equations that each obey are adjoined to the model equations. When command inputs and disturbance inputs were considered simply as functions of time, solutions of the servomechanism problem and the problem of accommodating disturbances were both termed "unrealizable" owing to their dependence on future values of these inputs, Reference 37, page 798.

The above approach presumes that the disturbances do not directly contribute to the measurements. This limits its use in the accommodation of disturbing effects which are not independent of the state

variables already in the model, such as truncated modes (which can affect the measurements directly).

3.6.2 Stochastic Disturbance Characterizations

In the absence of better data we might simply describe the disturbance $w(t)$ in (3.10) as white noise and take a stochastic approach to the problem. Stochastic approaches strive for good (optimal) performance in an average sense. A particular experiment might fail to meet requirements even without model errors. The expense of the experiment places a proportionate degree of responsibility on the control design(er). But, obtaining reliable statistical data for the system is often difficult. This fact, together with the fact that the theory for linear Gauss-Markov models is so well developed (and with a lot of encouragement from the central limit theorem, Reference 38, p. 96), leads to widespread use of the linear quadratic Gaussian (LQG) approach (see Reference 39 for an entire journal issue devoted to this subject). Three problems of the LQG approach which are of concern in the present research are: 1) Pursuing "long term" goals (the probabilistic approach) can lead to dire consequences in the "short term." (Successive short term "failures" can lead to system instability and Kalman filter divergence (denying long term goals); 2) The assumption of Gaussian distributions can be restrictive; 3) The assumption of white noise can cause problems (section 2.2 discussed some effects of correlated disturbances which are modeled as white).

3.7 Adaptive Approaches

The basic question motivating most adaptive approaches is "Are there better values for the _parameters_ (within the fixed model structure assumed) than those which were chosen a priori?" This question applies equally to deterministic and stochastic models. In Figure 2-1 and previous discussions we referred to parameters (A^j, B^j, C^j, M^j), $\hat{K}$, and G. It is these parameters that are available for adjustment in either the deterministic or stochastic case, although in the stochastic case $\hat{K}$ is usually changed via changes in $\hat{Q}$ through the $\hat{K}(\hat{Q})$ relation shown in Figure 2-2. The various algorithms employed to change the parameters will not be discussed here, but we are interested in the question "Does there exist _any_ set of parameters, within the model structure assumed, which will stabilize (or make acceptable) the controlled system?".

3.7.1 Deterministic Parameter Identification

In deterministic problems the model usually used for parameter adaptations is a "minimal realization" Reference 40, p.94, (a controllable, observable model). The limitation faced by such a scheme relates to the following arguments:

a. It is reasoned in this report that physical dynamical systems are _not_ completely controllable. (Supporting arguments will be offered).

b. Input/output relationships obtained in a controlled experiment (i.e., all initial conditions are zero) can lead one to select "minimal realizations" to describe the system. However, in actual operation of the system in its natural environment the uncontrollable

37

modes (which are always present) can be excited and there may not exist any set of parameters for the minimal realization which will yield acceptable closed loop operation. In the model reference adaptive control schemes, the objective is to force the physical system (a very large order "model") to behave as a low order system. Parameter adjustments are sometimes used to find the best low order approximation to the (high order) physical system, and usually complete controllability is presumed, as noted above.

3.7.2 Adaptive Kalman Filtering

Among the approaches to prevent Kalman filter divergence the most common theme is to prevent the Kalman gain from reaching its steady state value (as determined by the model, see Figure 2-2, thusly, "instructing" the filter to discredit the model. One way to do this is to artificially, or otherwise, increase the intensity of the plant noise (increase the covariance matrix $\hat{Q}$ in Figure 2-2, Reference 17-20). Most $\hat{Q}$ selection procedures are based upon:

- comparison of predicted with actual statistics of the measurement residuals, References 17 and 18.

- a priori experimentation, Reference 19,

- sensitivity approaches, References 41, 42, and 43.

Increasing $\hat{Q}$ can sometimes partially compensate for unmodeled correlated disturbances, although estimates of states of a stable system may also get worse with increasing $\hat{Q}$, Reference 22. Steady state gain prevention may also be accomplished by measurement "aging" (viewing the older measurements with less confidence). Exponential aging is proposed in Reference 44, and linear aging is used in Reference 45.

38

Jazwinski[46] and Crump[41] use a limited memory filter to serve the same purpose.

Finally, we are left with a basic requirement for the successful application of these adaptive techniques; that the model assumed possesses an "appropriate" structure.* Thus, errors in structure (that is, the order of the assumed model is too low) can be more critical than errors in parameters because the former errors may not be compensated even by adaptive techniques. Likewise, in the LQG problem, when correlated disturbances have been modeled as white noise there may not exist any values of the system model and covariance parameters $(A,B,C,M,\hat{Q},\hat{R},P_o)$ which will yield a stable system even with adaptive techniques, Reference 22. Correlated disturbances (either $w(t)$ or $v(t)$ in Figure 2-2) may, of course, be modeled by a Gauss-Markov process, Reference 18, adding order to the system model, but it is not clear how to select such a model. Astrom[1] has also pointed out that increasing the order of the model can compensate for parameter errors, and we have above discussed the fact that the converse is not true (changing parameter values may not compensate for errors in model order). Bryson and Henrikson[47] and Stubberud and Stear[48] accommodate correlated disturbances in the measurement process ($v(t)$ is correlated in Figure 2-2), by differentiating (in continuous models) or differencing (in discrete models) the measurements to make new "measurements" white. Differentiation can create practical problems

---

* We must keep in mind the fact that there are no well developed theories to help one determine what is an "appropriate" model for the particular control task at hand.

because of noise amplification. A more serious handicap with this approach, however, is due to the fact that the system is more sensitive to disturbance modeling errors than is the augmentation approach using Gauss-Markov models. This notion will be made more precise in the sequel.

4.0 Control by Model Error Estimation

In the discussion of the last section it was pointed out that the structure (order) of the model was very important to the success of the control objectives. Several of the schemes discussed were concerned directly with changing the model order and we will pause briefly to show the relationship of these methods with the one to be proposed in this section.

Refer to Figure 1-1 where the proposal was made that models be constructed in two steps.[*] The first step $(\mathcal{M}_1 \rightarrow \mathcal{M}_3)$ was a model reduction (or in the language of the dynamicist, mode truncation) which left us with a model $\mathcal{M}_3$ which might be quite sensitive to the model errors associated with truncated modes, neglected external disturbances, and parameter errors. The second step of the proposed modeling plan of Figure 1-1 requires the expansion of the model (but not up to the order of $\mathcal{M}_1$) with the hope that the added modes, whose description is to be accomplished in this section, would lend to the composite system a degree of model error insensitivity. Now to be more precise. The singular perturbation approaches of Kokotovic[7] and others, provide a way to obtain the approximate solution for the higher order system controller based upon considerations of a lower order model. (Specifically, the method yields, for the linear regulator problem, an approximation of the larger dimensional, $n_1 \times n_1$ Riccati matrix based upon asymptotic expansions of the Riccati matrix in terms of a small parameter ϵ. When $\epsilon = 0$ the state equations reduce from order n_1 to order n_3.) This then, provides one method to obtain an $\mathcal{M}_2$

_____
[*]That is, two steps after the construction of the conservative model $\mathcal{M}_1$.

model: consider the model from which $\mathscr{A}_3$ is a truncation ($\mathscr{A}_1$) and implement an approximation of the higher order controller. Care must be taken to restrict the model $\mathscr{A}_1$ to an order which can be implemented (in the state estimator of the controller); usually this restriction requires a low order for controllers useful for spacecraft, for instance. The other methods to be discussed, including the one outlined in this research, <u>begin</u> with the smaller model $\mathscr{A}_3$, although the presumption is made that $\mathscr{A}_3$ represents a truncation of a higher order model, $\mathscr{A}_1$ (even though we might never actually construct $\mathscr{A}_1$) and is attended with the associated model errors.

The sensitivity methods of Cruz, Perkins[0] and others provide another way to construct a model $\mathscr{A}_2$ from the model $\mathscr{A}_3$, and that is to augment to $\mathscr{A}_3$ the parameter sensitivity subsystem (the $\dot{s} = f(p,x,u,s)$ equations in (3.2b)). This sensitivity subsystem is generally uncontollable, as is the "model error system" to be proposed in this section. However, a key difference between the methods is that the sensitivity variables $s(t)$ are desired to be minimized (to be penalized in the cost functional of (3.1)) and their uncontrollability therefore creates difficulties. The variables of the "error systems" described below do not appear in the cost functional and are only added to the system to make the estimates of the state variables (of $\mathscr{A}_3$) more reliable. Therefore their uncontrollability creates fewer problems.

One further disadvantage of the sensitivity method of constructing $\mathscr{A}_2$ is to be noted. This approach expands the model from $\mathscr{A}_3$ to $\mathscr{A}_2$ based solely upon information contained in $\mathscr{A}_3$, the truncated model. As mentioned in Section 3.0, this disadvantage is not shared by the

singular perturbation methods[7] or the method proposed in this report, both of which "consult" $\mathcal{S}_1$ before constructing $\mathcal{S}_2$. (Actually, $\mathcal{S}_1 \equiv \mathcal{S}_2$ in the singular perturbation method.)

There are numerous papers which concern themselves with feedback through dynamical compensation, and as such can be considered as methods to obtain an augmented system $\mathcal{S}_2$. Johnson[2] and Johnson and Skelton[3] provide a controller which is based upon the construction of an observer to estimate the states of the equations that the external disturbance is assumed to obey. Thus, in that paper $\mathcal{S}_2$ is obtained by augmenting to $\mathcal{S}_3$ the equations of the external disturbance. Johnson labels such controllers "disturbance accommodating".[4] Davison[6] also accommodates external disturbances and command inputs which satisfy the (same) given differential equations. Bhattacharyya and Pearson,[49] include both equations for external disturbance and input command signals and convert the resulting servomechanism problem (via the transformation to an "error system")* into a linear regulation problem which is then solved. Algebraic conditions are given for the existence of a stabilizing solution assuming availability of all states of the composite system (i.e. perfect state estimation). Some sufficiency conditions are also established for the existence of a stabilizing output-feedback controller. A conclusion of Reference 49 is that the model $\mathcal{S}_2$ (which is now the composite of $\mathcal{S}_3$, the disturbance equations, and the input command equations) is not always stabilizable by state

---

* Bhattacharyya and Pearson's "error systems" describes quantities which the control designer wishes to drive to zero. The model "error system" of this report, which will shortly be defined, is not the same.

(of $\mathcal{A}_2$) feedback. The present study <u>begins</u> with the condition of uncontrollable and even unstabilizable models $\mathcal{A}_2$, and argues that a practical concern is the stabilization of a model different from $\mathcal{A}_2$. In the sequel the model $\mathcal{A}_2$ is used only for determining the desired structure of the dynamical controller (control law plus state estimator). Also the papers described above say very little about how one should go about constructing the models which are augmented to $\mathcal{A}_3$. It is a primary concern of this report to provide a procedure for determining the parameters of the model $\mathcal{A}_2$, given the model $\mathcal{A}_3$. This emphasis makes this study primarily one of modeling methods rather than one of new control theory development, as noted previously; however the inherent interdependence of the modeling and control problems obscures this distinction. Now, let us proceed to develop the "model error system" which we will augment to $\mathcal{A}_3$ to form $\mathcal{A}_2$. Firstly we must provide suitable definitions of the model error.

4.1 Model Error Definitions

In this research the letter z always denotes the measurement vector, and the letter y always denotes the output quantities which are desired to be regulated by control. In other words, for the optimal control problems we will discuss, y^j appears as those variables associated with model $\mathcal{A}_j$ which we penalize in the performance functional

$$J[u^0] = \frac{1}{2} y^j(T)^T H \, y^j(T) + \frac{1}{2} \int_0^T (y^{jT} Q \, y^j + u^{0T} R \, u^0) \, dt$$

$$y^j \in R^k$$

$$0 \le k \le n_j$$

where Q, R are always positive definite, and H is positive semidefinite. The variables $z_i^j(t)$ $i = 1 - \ell$ are those variables which are actually measured (as nearly as the model $\mathscr{A}_j$ can predict). The vector $y^j \in R^k$ where $0 \leq k \leq n_j$. Thus, one can sometimes choose $k = n_j$ if, he wishes to obtain that extra quality of smoothness in optimal trajectories $x(t)$ afforded by the least squares theory when all of the n_j states of model $\mathscr{A}_j$ are penalized. The superscript j means that the variable is associated with the model $\mathscr{A}_j$. The notation $u^o(t)$ denotes the control actually input to the physical system. The notation $z^o(t)$ refers to that time record of measurements actually obtained from the physical system.

Consider the model $\mathscr{A}_j$

$$
\mathscr{A}_j \left\{
\begin{array}{ll}
\dot{x}^j = A^j x^j + B^j u^o & x^j \in R^{n_j} \\
\begin{pmatrix} y^j \\ z^j \end{pmatrix} = \begin{bmatrix} C^j \\ M^j \end{bmatrix} x^j &
\begin{array}{l} u^o \in R^m \\ y^i \in R^k \\ z^j \in R^\ell \end{array}
\end{array}
\right.
\tag{4.1}
$$

and the model $\mathscr{A}_i$

$$
\mathscr{A}_i \left\{
\begin{array}{ll}
\dot{x}^i = A^i x^i + B^i u^o + w^i & x^i \in R^{n_i} \\
\begin{pmatrix} y^i \\ z^i \end{pmatrix} = \begin{bmatrix} C^i \\ M^u \end{bmatrix} x^i &
\begin{array}{l} y^o \in R^m \\ y^i \in R^k \\ z^i \in R^\ell \end{array}
\end{array}
\right.
\tag{4.2}
$$

Suppose $n_i > n_j$ and consider a particular similarity transformation on $\mathscr{A}_i$ such that

$$
x^i = T \cancel{X}
\tag{4.3}
$$

$$
T^{-1} A^i T = \begin{bmatrix} A^j & A_{12} \\ A_{21} & A_{22} \end{bmatrix}
\tag{4.4}
$$

45

$$T^{-1}B^i = \begin{bmatrix} B^j \\ B_2 \end{bmatrix} \tag{4.5}$$

$$M^i T = [M^j 0] \tag{4.6}$$

$$C^i T = [C^j 0] \tag{4.7}$$

where all matrices without the j superscript are arbitrary. Note that by counting the constraint equations,

$$[T^{-1}A^i T]_{11} = A^j \qquad (n_j^2 \text{ eqs.}) \tag{4.8}$$

$$[T^{-1}B^i]_1 = B^j \qquad (n_j m \text{ eqs.}) \tag{4.9}$$

$$M^i T = [M^j, 0] \qquad (n_i \ell \text{ eqs.}) \tag{4.10}$$

$$C^i T = [C^j, 0] \qquad (n_i k \text{ eqs.}) \tag{4.11}$$

$$T^{-1} T = I \qquad (\tfrac{1}{2} n_i (n_i + 1) \text{ eqs.}) \tag{4.12}$$

A necessary condition for the existence of such n_i^2 unknown of T is obtained[*], requiring the model $\mathscr{A}_i$ to be sufficiently larger than the model $\mathscr{A}_j$ such that

$$n_i \geq \frac{2(k+\ell)+1}{2} \left[1 + \sqrt{1 + \frac{8\, n_j(n_i+m)}{(2(k+\ell)+1)^2}} \right] \tag{4.13}$$

It is helpful in gaining insight in what follows to assume that such a T exists although the main results do not depend upon it. In fact, such a T need not ever be computed to utilize the procedures suggested in the summary of Section 5.0. Note also that if the constraints on T are relaxed by ignoring (4.11), then, (4.13) is relaxed to

$$n_i \geq \frac{2\ell+1}{2} \left[1 + \sqrt{1 + \frac{8\, n_j(n_j+m)}{(2\ell+1)^2}} \right] \tag{4.14}$$

---

[*]The requirement $n_i^2 \geq (2(k+\ell)+1)n_i + 2\,n_j(n_j+m)$ yields at least as many unknowns as constraint equations.

46

and by ignoring, instead, the constraint (4.10), (4.13) is relaxed to

$$n_i \geq \frac{2k+1}{2} \left[1 + \sqrt{1 + \frac{8n_j(n_j+m)}{(2k+1)^2}} \right] \tag{4.15}$$

Prompted by these observations we offer the following definitions.

Definition I: If there exists a vector, $e_z^{ji}(t,x_o^i,x_o^j)$ such that $z^i(t) \equiv z^j(t)$, where $z^i(t)$ obeys (4.2) and $z^j(t)$ obeys

$$\begin{aligned}
\dot{x}^j &= A^j x^j + B^j u^o + e_z^{ji}(t) \\
z^j &= M^j x^j
\end{aligned} \tag{4.16}$$

then such a vector is called a "model error vector with respect to measurement".

Definition II: If there exists a vector, $e_y^{ji}(t,x_o^i,x_o^j)$, such that $y^i(t) \equiv y^j(t)$, where $y^i(t)$ evolves from (4.2) and $y^j(t)$ evolves from the system

$$\begin{aligned}
\dot{x}^j &= A^j x^j + B^j u^o + e_y^{ji}(t) \\
y^j &= C^j x^j
\end{aligned} \tag{4.17}$$

then such a vector is called a "model error vector with respect to output".

Definition III: If there exists a vector $e^{ji}(t,x_o^i, x_o^j)$ such that

$$\begin{pmatrix} y^i(t) \\ z^i(t) \end{pmatrix} \equiv \begin{pmatrix} y^j(t) \\ z^j(t) \end{pmatrix} \tag{4.18}$$

where $\begin{pmatrix} y^i(t) \\ z^i(t) \end{pmatrix}$ evolves from (4.2) and $\begin{pmatrix} y^j(t) \\ z^j(t) \end{pmatrix}$ evolves from

$$\begin{aligned}
\dot{x}^j &= A^j x^j + B^j u^o + e^{ji}(t) \\
\begin{pmatrix} y^j \\ z^j \end{pmatrix} &= \begin{bmatrix} C^j \\ M^j \end{bmatrix} x^j
\end{aligned} \tag{4.19}$$

then such a vector is called a "model error vector with respect to output and measurement". (In the sequel it is this model error vector, e^{ji}, which finds the most frequent use and it will be convenient to shorten the name to "model error vector".)

The first observation one can make concerning these definitions is that if no restrictions* are placed upon the functions of time, $e^{ji}(t)$, $e_y^{ji}(t)$, $e_z^{ji}(t)$, then all such model error vectors defined above exist, assuming observability of the model $\mathscr{A}_j$ both in output and measurement (the pairs (A^j, M^j) and (A^j, C^j) are observable). To show this, consider only the definition of $e_y^{ji}(t)$. The output of $\mathscr{A}_i$, differenced from the output of $\mathscr{A}_j$, is given by

$$y^i(t) - y^j(t) = C^i \Phi^i(t,o) x_o^i - C^j \Phi^j(t,o) x_o^j$$

$$+ \int_o^t \left\{ C^i \Phi^i(t,\sigma)(B^i u^o + w^i) - C^j \Phi^j(t,\sigma)(B^j u^o + e_y^{ji}) \right\} d\sigma \qquad (4.20)$$

The requirement $y^i(t) \equiv y^j(t)$ is satisfied by the choice**

$$e_y^{ji}(\sigma) = [C^j \Phi^j(t,\sigma)]^T \left\{ [C^i \Phi^i(t,o) x_o^i - C^j \Phi^j(t,o) x_o^j] \delta(t-\sigma) \right.$$

$$\left. + C^i \Phi^i(t,\sigma) w^i + [C^i \Phi^i(t,\sigma) B^i - C^j \Phi^j(t,\sigma) B^j] u^o(\sigma) \right\} \qquad (4.21)$$

where $\delta(t-\sigma)$ is the Dirac delta function and T denotes the generalized psuedo-inverse of Penrose. The matrices Φ^i and Φ^j denote the state

---

*That is, if the functions are allowed to contain impulses, doublets, etc.

**Note that $C^j \Phi^j(t,\sigma) e^{ji}(\sigma) \equiv 0$ if and only if $e^{ji}(\sigma) \equiv 0$ under the assumption (actually the definition) of observability of the model $\mathscr{A}_j$.

transition matrices for A^i and A^j, respectively. Note also that the above definitions of model error are relative to two models (of the physical system) and thus have nothing to do with the physical systems itself.

Now let us view the error vector, $e^{ji}(t, x_o^j, x_j^i)$, from another perspective. Consider the transformed system $\mathscr{A}_i$ in (4.2)-(4.7) and see that the vector which can be added to the plant equations of model $\mathscr{A}_j$ to make its outputs identically equal to those of the larger (and presumed better) model $\mathscr{A}_i$ is given by the "error system"

$$e^{ji}(x_o^j, x_o^i, t) = A_{12}x^t + \mathscr{E}_1 w^1$$

$$\dot{x}^t = A_{22}x^t + A_{21}x^j + B^t u^o + \mathscr{E}_2 w^1 \quad , \quad \begin{bmatrix} \mathscr{E}_1 \\ \mathscr{E}_2 \end{bmatrix} \triangleq T^{-1} \quad , \quad \begin{bmatrix} B^j \\ B^t \end{bmatrix} \triangleq T^{-1} B^i \quad (4.22)$$

where the transformed n_i-vector, $\mathcal{X}$, in (4.3) has been partitioned with the designation

$$\mathcal{X} = \begin{pmatrix} x^j \\ x^t \end{pmatrix} \qquad \begin{array}{l} x^j \in R^{n_j} \\ x^t \in R^{n_i - n_j} \end{array} \qquad (4.23)$$

To see this simply write the $\mathscr{A}_i$ model in (4.2) in the transformed coordinates (4.3)-(4.7) to get (writing the separate equations for the partitioned parts x^j and x^t).

$$\mathscr{A}_i \begin{cases} \begin{cases} \dot{x}^j = A^j x^j + B^j u^o + e^{ji}(t) \\ \begin{pmatrix} y^j \\ z^j \end{pmatrix} = \begin{bmatrix} C^j \\ M^j \end{bmatrix} x^j \end{cases} & (4.24a) \\ \\ \begin{cases} \dot{x}^t = A_{22}x^t + A_{21}x^j + B^t u^o + \mathscr{E}_2 w' \\ e^{ji}(t) = A_{12}x^t + \mathscr{E}_1 w' \end{cases} & (4.24b) \end{cases}$$

and compare (4.24a) with (4.19) and (4.24b) with (4.22). We have

presumed also that the initial conditions, x_o^j, associated with $\mathcal{A}_j$ in

(4.1) have also been applied to the n_j-vector x^j in (4.24a).

If we wish to make reference to models which approximate the

physical system (in an arbitrarily small region around the null solu-

tion of the (linear) model equations) well enough to serve as an

evaluation* model then such models will necessarily be of high order

so that the condition (4.13) is satisfied and the interpretation of the

model error vector in (4.19) is useful. Use of the transformation

defined in (4.3) clearly focuses the fact that, under the conditions

discussed above, there exists a model error vector which can be added

to any linear model $\mathcal{A}_j$ which accommodates the effects of neglected

external disturbances, $w(t)$, truncated modes, corresponding to the

truncated states $x^t(t)$ and parameter errors. To appreciate the

inclusion of parameter errors in this statement, note that only the

parameters of the transformation T would change to give $A^j + \Delta A^j$,

$B^j + \Delta B^j$, on the right hand side of (4.4)-(4.7), changing the second

equation of (4.24b) to

$$e^{ji}(t) = \Delta A^j x^j + \Delta B^j u^o + A_{22} x^t + \mathcal{E}_1 w' \qquad (4.25)$$

Parameter errors in C^j and M^j are accommodated in the output and

measurement residuals to be defined in what follows.

---

*An "evaluation" model is one which represents the physical system well
enough to be used to evaluate candidate control schemes via analysis
or simulation by computer in lieu of physical system testing.

Any approximations of the model error vector which we will subsequently use will have residual errors associated with them, requiring still further definitions. Suppose an approximation of $e^{ji}(t)$ is written $\hat{e}^{ji}(t)$ (when, in the sequel, it is clear which two models $\mathscr{A}_i$, and $\mathscr{A}_j$ are being referenced the j,i superscripts will be omitted), then define a residual model error

$$\tilde{e}^{ji} \triangleq e^{ji} - \hat{e}^{ji} \tag{4.26}$$

Now the relative residual errors in the outputs and measurements are defined as a <u>consequence</u> of the residual in the model error and are defined by

$$\begin{pmatrix} \tilde{y}^{ji}(t) \\ \tilde{z}^{ji}(t) \end{pmatrix} \triangleq \begin{bmatrix} C^j \\ M^j \end{bmatrix} \int_o^t \Phi^j(t,\sigma)\ \tilde{e}^{ji}(\sigma)d\sigma \tag{4.27}$$

Now it follows that

$$\begin{pmatrix} y^i \\ z^i \end{pmatrix} = \begin{pmatrix} y^j \\ z^j \end{pmatrix} + \begin{pmatrix} \tilde{y}^{ji} \\ \tilde{z}^{ji} \end{pmatrix} \tag{4.28}$$

Thus the relative output and measurement residuals are defined by (4.27) if one, in turn, recalls the definition of $\tilde{e}^{ji}$ and e^{ji}. Equivalently, and more conveniently, one can take (4.28) as the definition of the relative output and measurement residuals, $\tilde{y}^{ji}$, $\tilde{z}^{ji}$. Finally, the measurements evolve from the physical system and not from the model of the physical system. Define any difference by

$$\begin{pmatrix} \tilde{y}^i \\ \tilde{z}^i \end{pmatrix} \triangleq \begin{pmatrix} y^o - y^i \\ z^o - z^i \end{pmatrix} = \begin{pmatrix} y^o - C^i x^i \\ z^o - M^i x^i \end{pmatrix} \tag{4.29}$$

51

The error relative to the physical system (called the output or measurement residual) and the error relative to two models (called the relative output or measurement residual) can be related by, using (4.28), (4.29)

$$\begin{pmatrix} \tilde{y}^i \\ \tilde{z}^i \end{pmatrix} = \begin{pmatrix} y^o \\ z^o \end{pmatrix} - \begin{pmatrix} y^j \\ z^j \end{pmatrix} - \begin{pmatrix} \tilde{y}^{ji} \\ \tilde{z}^{ji} \end{pmatrix}$$

or by the "chain rule"

$$\begin{pmatrix} y^o \\ z^o \end{pmatrix} = \begin{pmatrix} y^j \\ z^j \end{pmatrix} + \begin{pmatrix} \tilde{y}^{ji} \\ \tilde{z}^{ji} \end{pmatrix} + \begin{pmatrix} \tilde{y}^i \\ \tilde{z}^i \end{pmatrix} \qquad (4.30)$$

It should be noted that if the models of this section were restricted to the unforced case ($w'(t) \equiv 0$ in (4.2), and $e^{ji}(t) \equiv 0$ in (4.19)) then the model matching results of Reference 50 can be applied, to wit:

Lemma 1: (Moore and Silverman[50], p. 492)

Refer to models $\mathscr{A}_j$ in (4.1) and $\mathscr{A}_i$ in (4.2). Let

$$m \leq k$$

and define

$$\delta_o^i \triangleq \text{rank} \left[B^i, A^i B^i, \ldots \ldots A^{i^{(n_i-1)}} B^i \right]$$

$$\delta_o^i \triangleq \text{rank} \left[(C^i)^T, (C^i A^i)^T, \ldots \ldots \left(C^i A^{i^{(n_i-1)}} \right)^T \right]$$

$$\delta_\rho^j \triangleq \text{rank} \left[B^j, A^j B^j, \ldots \ldots \ldots A^{j^{(n_j-1)}} B^j \right]$$

$$\delta_o^j \triangleq \text{rank} \left[(C^i)^T, (C^j A^j)^T, \ldots \ldots \ldots \left(C^j A^{j^{(n_j-1)}} \right)^T \right]$$

Then the model $\mathscr{A}_j$ is "zero state equivalent in the output" to $\mathscr{A}_i$ if and only if

$$c^j(A^j)^{\alpha} B^j = c^i(A^i)^{\alpha} B^i \qquad\qquad (4.31)$$

for $\qquad \alpha = 0, 1, \ldots \nu$ where

$$\nu \doteq \max\left\{\delta_c^1 + \delta_o^1 - 1, \quad \delta_c^j + \delta_o^j - 1\right\} \quad.$$

A corresponding result of "zero state equivalence in the measurements" is available by replacing C^i by M^i everywhere in the above lemma. "Zero state equivalence in the output (measurement)" means $y^j(t) \equiv y^i(t)$, (or, $z^j(t) \equiv z^i(t)$), if the initial conditions, $x^j(o)$ and $x^i(o)$, are all zero. In the physical system it is not usually possible to keep the initial conditions (which correspond to all the variables x_k^i, $k=1, -m_i$) zero, especially for large n_i. For instance, in spacecraft the structural modes of vibration are unavoidably excited by noisy actuators (driven by noisy sensor measurements), and other disturbances. Moreover, some of these modes (there are actually an infinite number of them in the physical structure) will correspond to those truncated in the reduced model $\mathcal{A}_j$. Therefore, input/output relationships of model $\mathcal{A}_j$ cannot account for $x^t(o) \neq 0$ (see (4.23)). Also, input/output relationships cannot account for uncontrollable states which are reasoned to be present in all physical dynamical systems. It is for these reasons that input/output procedures will not be pursued further here.

In the following section we wish to consider a description of the synthetic modes assumed for the error system.

4.2 Selection of the Synthetic Modes for the Model Error System

Recall that the exact control problem which we would like to solve, neglecting implementation constraints, is the minimization of

$$J(u^o) = \frac{1}{2} \int_o^T (y^{1^T} Q\, y^1 + u^{o^T} R\, u^o)\, dt + \frac{1}{2}\, y^1(T)^T H\, y^1(T) \tag{4.32}$$

subject to the very large model $\mathcal{A}_1$

$$\mathcal{A}_1 \begin{cases} \dot{x}^1 = A^1 x^1 + B^1 u^o + w^1 & x^1 \in R^{n_1} \\[2mm] \begin{pmatrix} y^1 \\ z^1 \end{pmatrix} = \begin{bmatrix} C^1 \\ M^1 \end{bmatrix} x^1, \quad x^1(o) = x_o^1 & \begin{array}{l} y^1 \in R^k \\[1mm] z^1 \in R^{\ell} \end{array} \end{cases} \tag{4.33}$$

where the coordinates of (4.23), (4.24) are presumed. Now (4.32), (4.33) can be considered the exact variational problem which we can only solve by _approximate_ solution methods, owing to its large dimension. Alternately, this problem can be replaced by an _approximate_ problem which we can solve _exactly_.

The approximate problem we might solve is as follows: Consider the decomposition of (4.33) as given in (4.3)-(4.7), (4.24) but instead of utilizing the exact error system (4.24b) to generate the vector $e^{31}(t)$, "curve fit" $e^{31}(t)^*$ with a set of d approximating functions $f_i(t)$, $i = 1,2,\ldots d$.

$$\hat{e}^{31}(t) = \sum_{i=1}^{d} p^i f_i(t) = P^o f \qquad \begin{array}{l} \hat{e}^{ji} \in R^{n_j} \\[2mm] f \in R^d \end{array} \tag{4.34}$$

We must pay particular attention to the initial conditions used in (4.33) which are used to generate the particular $e^{31}(t)$ we wish to approximate, as discussed in the next section.

---

* The smaller model is labeled $\mathcal{A}_3$ and the larger model in this discussion is $\mathcal{A}_1$ see Figure 1-1. Hence, through this section e^{ji} becomes e^{31}.

54

4.2.1 Model Error Compensation by Open Loop Signals

In this section (4.2.1) we discuss a technique of model error compensation which we will not later use and this section only provides a basis for comparison with the method (in the following section, 4.2.2) which we do intend to incorporate into our overall design scheme.

For a given $e^{31}(t)^*$ and a given set of functions, $f_i(t)$, the coefficients P of (4.34) can be chosen to minimize the least squares of the error, $e^{31}(t) - \hat{e}^{31}(t) \stackrel{\Delta}{=} e$.

$$J(P^O) = \int_O^T e(t)^T g(t) \, e(t) dt \qquad (4.35)$$

or the least squares of error plus error rate

$$J(P^O) = \int_O^T (e^T \, \dot{e}^T) Q_e(t) \begin{pmatrix} e \\ \dot{e} \end{pmatrix} dt \qquad (4.36)$$

where $g(t)$ is a weighting scalar, $Q_e(t)$ is a positive definite matrix $\forall \, t \in (\Delta, T)$ and where T is the entire interval of the control problem. (see (4.32)). A shorter interval for consideration in (4.35) will be discussed in Section 4.3.

The use of least squares methods must be carefully considered so as not to allow large instantaneous errors (4.36) (and even sustained oscillations of large magnitude). The criteria will enforce a certain

---

*The presumption of the availability of $e^{31}(t)$ has a serious consequence. Nothing has been said so far which relieves us of the duty to actually construct the larger model $\mathcal{A}_1$, select a truncation (a reduced model) and generate e^{31} according to (4.24b). Some relief from the press of this requirement is found in the next section. For a guess for $e^{31}(t)$, one could "assume some modes" which are believed to be present in (4.24b), such as, for example, some structural modes of vibration which have not been included in the truncated model $\mathcal{A}_3$.

smoothing quality by simultaneously keeping error rates and errors

small. Alternately, the approximating functions, $f_i(t)$, of (4.34) may

themselves be selected so that even (4.35) yields certain smoothing

qualities, and this is a subject of Section 5.0. For the present, we

will consider (4.35) for the determination of P, noting that (4.36) may

be a wiser choice for arbitrary selections of approximating functions,

$f_i(t)$. The P which minimizes (4.35) is given by (see Appendix A)

$$P^o = \int_0^T g(t) e^{31}(t) f^T(t) dt \left[\int_0^T g(t) f(t) f^T(t) dt \right]^{-1} \quad (4.37)$$

Therefore, one approximation of the original problem (4.32), (4.33) is:

Minimize,

$$J(u^o) = \frac{1}{2} \int_0^T \left(y^{2^T} Q y^2 + u^{o^T} R u^o \right) dt + \frac{1}{2} y^2(T)^T H \, y^2(T)$$

subject to*

$$\mathcal{A}_2 \left\{ \begin{array}{l} \dot{x}^2 = A^3 x^2 + B^3 u^o + P^o f \\[2mm] \begin{pmatrix} y^2 \\ z^2 \end{pmatrix} = \begin{bmatrix} C^3 \\ M^3 \end{bmatrix} x^2 \end{array} \right. \quad (4.38)$$

where P^o is specified by (4.37).

The solution of this type of control problem (with the forced

plant equation this was originally called the servomechanism problem)

can be found in a number of places. For constancy of notation see p. 15

of Reference 3.

---

*When an error vector, $\hat{e}^{31}$, has been added to the (truncated) model $\mathcal{A}_3$,
  the new model will be labeled $\mathcal{A}_2$, even though some parameters of $\mathcal{A}_2$
  are retained from $\mathcal{A}_3$, see (4.38).

56

$$u^o = - R^{-1} B^{3^T} (Kx^2 + h(t))$$

where h, K satisfy   (4.39)

$$\dot{K} = - KA^3 - A^{3^T}K + KB^3 R^{-1}B^{3^T}K - C^{3^T}QC^3 \qquad K(\tau) = H$$

$$\dot{h} = \left[KB^3 R^{-1}B^{3^T} - A^{3^T} \right] h - KPf \qquad h(\tau) = 0$$

where it is assumed that the pair (A^3, C^3) is observable and the pair $(A^3 B^3)$ is controllable. The difficulties with this approximation of problem (4.32), (4.33) include:

1. The error vector, $e^{31}(t)$, which includes truncated states and external disturbances must be reliably predicted over the <u>entire</u> control interval. (If $e^{31}(t)$ is considered as only an external disturbance then periodic updates in P and f may be helpful.) Thus, $e^{31}(t)$ can be quite sensitive to the system (and model $\mathcal{M}_1$) initial conditions, see (4.24b).

2. The control problem for the larger model, $\mathcal{M}_1$, may have to be solved to obtain a suitable $e^{31}(t)$. (See Rodgers and Sworder[11] for a method to obtain a smaller model by solving the control problem for both the small and large models and minimizing the difference.)

We would welcome now a scheme for computing $\hat{e}^{31}(t)$ without having to predict $e^{31}(t)$ accurately for all time, $t \in [0,T]$. We wish, for example, to allow for changes in parameter values from time to time, changes in the nature of external disturbances, and changes in the initial conditions of truncated modes (see (4.25) for the way in which these affect the actual model error vector, $e^{31}(t)$). The a priori specification of $\hat{e}^{31}(t)$ $\forall t \in [0,T]$, as presented in the "open loop" method above, precludes such flexibility. The fundamental difference

in the approach of the next section is the recognition of the fact that, as far as the construction of a state estimator is concerned, prediction of model errors only over a <u>short</u> time into the future is of interest. This fact allows us the freedom of modeling only those modes in $e^{31}(t)$ which <u>appear</u> to be present in a preselected (short) interval of time, $\tau \ll T$. Naturally, it follows that the choice for τ is an important parameter in the model. By relating τ to certain properties of the control problem a (modest) mathematical relationship is established between the modeling problem and the control problem.

Now soon we will restrict our attention to the class of functions, $e^{ji}(t)$, which are d^{th} differentiable, because we wish to approximate $e^{ji}(t)$ as the output of a dth order differential equation. Hence the existence of $e^{ji}(t)$ which was guaranteed for unrestricted functions of time, does not carry over to the functions $\hat{e}^{ji}(t)$ which are restricted (to be differentiable). However, in many practical situations the physical system is sufficiently well behaved to consider only (linear) differential equation models of all effects. In the following section, however, we do not yet restrict ourselves to differential equations to describe the synthetic "modes" assigned to the error system.

4.2.2 Model Error Compensation by Closed Loop Signals

In this section we propose an approximation of the model error vector

$$\hat{e}^{31} = Pf \tag{4.40}$$

which allows the coefficients, P_{ji}, of the preselected function $f_i(t)$, $i = 1,2,\ldots d$, to change with time (P must change slowly compared to a preselected time interval, τ, where τ will soon be defined). The

58

matrix P is defined similarly to (4.37) except that T is replaced by τ

$$P = \int_0^\tau g(t) e^{31}(t) f^T(t) dt \left[\int_0^\tau g(t) f(t) f^T(t) dt \right]^{-1} \qquad (4.41)$$

Now construct a differential equation which f(t) obeys. Find a matrix D such that f(t) is an eigenfunction of the system,

$$\dot{\gamma} = D\gamma \qquad (4.42)$$

Now replace f(t) in (4.40) by a set of "synthetic" variables $\gamma(t) \in R^d$ and write for the error vector approximation the "error system"

$$\hat{e}^{31} = P\gamma(t) \qquad (4.43)$$

$$\dot{\gamma} = D\gamma \qquad (4.44)$$

where $\gamma(t)$ in (4.43) is <u>any</u> solution of (4.44). Then

$$\hat{e}^{31} = Pe^{D(t-t_o)} \gamma(t_o) \qquad (4.45)$$

where the eigenmodes of the matrix D represent the synthetic modes assumed to be present in $e^{31}(t)$, <u>over an interval τ</u>. This construction of the error system relieves us of the responsibility of predicting $e^{31}(t)$ for all future time, $t \in [0,T]$, as required in (4.37). The task of "assuming modes" which characterize $e^{31}(t)$ only over short intervals is a far simpler (realizable) task.

If $\gamma(t_o)$ (see (4.45)) can be updated in real time by state estimation of the error subsystem (4.43), (4.44) then the estimate $\hat{e}^{31}(t)$ will not be a function of time which is <u>fixed a priori</u> as in the case of Section 4.2. In this way another alternative approximation to problem (4.32), (4.33) is formed: Minimize

$$J(u^o) = \frac{1}{2} \int_0^T \left(y^{2^T} Q y^2 + u^{o^T} R u^o \right) dt + \frac{1}{2} y^2(T)^T H y^2(T) \qquad (4.46a)$$

subject to the model $\mathcal{A}_2$

$$\mathcal{A}_2 \begin{cases} \dot{x}^2 = A^2 x^2 + B^2 u^o \\ \begin{pmatrix} y^2 \\ z^2 \end{pmatrix} = \begin{bmatrix} C^2 \\ M^2 \end{bmatrix} x^2 \end{cases} \tag{4.46b}$$

where

$$x^2 \triangleq \begin{pmatrix} x^3 \\ \gamma \end{pmatrix} \qquad x^3: n_3- \text{ vector of physical variables}$$

$$A^2 = \begin{bmatrix} A^3 & P(\tau,d) \\ 0 & D(\tau,d) \end{bmatrix}, \quad B^2 = \begin{bmatrix} B^3 \\ 0 \end{bmatrix} \qquad \gamma: d - \text{ vector of synthetic (model error) variables}$$

$$C^2 = [C^3,0], \quad M^2 = [M^3,0] \qquad x^2: n_3 + d = n_2 - \text{ vector}$$

again observability of (A^3,C^3) and controllability of (A^3,B^3) are assumed and where $P(\tau,d)$ is found by (4.41)[*], and $\gamma(t)$ is to be estimated along with $x^3(t)$ in real time. The current model error vector estimate will be $\hat{e}^{31} = P\gamma(t)$. The solution to this problem is given by[**]

$$u^o = -R^{-1} B^{2^T} K^2 x^2 \tag{4.47}$$

where K^2 satisfies the matrix Riccati equation

$$\dot{K}^2 = -K^2 A^2 - A^2 K^2 + K^2 B^2 R^{-1} B^{2^T} K^2 - C^{2^T} Q C^2 \tag{4.48}$$

$$K^2(T) = C^{2^T} H C^2$$

---

[*] The order of the error system, d, and τ, an effective "observation window" are as yet unspecified.

[**] The initial condition $x^3(t_o) = x_o$ in (4.46) is presumed known and $\gamma(t_o)$ is actually unknown. However in the solution of the control problem, (4.47), $\gamma(t_o)$ is assumed known. An estimator for $x^3(t)$, $\gamma(t)$ will be constructed in Section 4.5 which will converge quickly compared to τ. The enforcing of this "separation principle" between the control and estimation problems must be done with care and more will be said about this in the sequel.

(In the next section a partitioned form of this equation will be examined.)

The addition of the "synthetic modes" (4.43), (4.44) to the model $\mathcal{A}_3$, (4.1) to form the model $\mathcal{A}_2$ (4.46) is similar in spirit to the "synthetic modes" Likins[66] uses in the modeling of dynamical structures to account for incorrect coupling forces between the rigid body and flexible appendages when the model has been truncated by some other criteria (such as keeping "lowest frequency" modes). The method also has some conceptual similarities with the "assumed modes" and Rayleigh Ritz procedures for modeling dynamical structures,[51] p. 253, although in the present work the assumption of mode "shapes" in (4.43) is in time rather than in space as in Rayleigh Ritz procedures. One choice for the matrix D, for instance, is to "assume" some eigenvalues λ_i, $i = 1,\ldots d$ of the matrix A_{22} in (4.22) and write $D = \begin{bmatrix} \lambda_1 & 0 \\ & \ddots & \\ 0 & & \lambda_d \end{bmatrix}$. The eigenfunctions of this matrix then form the definition of the approximating functions, f, in (4.40).

Even for a given dimension for n_2, such as might be imposed upon a state estimator synthesis, it is not always clear how many modes to associate with the physical variables x^3 and how many to allow as "synthetic modes" in $\gamma(t)$. The answer in a particular problem will depend upon the $f_i(t)$ chosen and the uncertainties associated with the model $\mathcal{A}_3$.

For example, suppose the object to be controlled is a free dynamical structure (such as a spacecraft) and there is some uncertainty in the (small) modal damping assigned to a particular mode in the vibration equations. Then if the frequency of vibration, ω, is reliably known

the mode in question might be truncated from the physical variables in the model (deleted from $\mathscr{A}_3$) and considered in the synthetic variables $\gamma_i(t)$ of the error system (4.43), (4.44) by choosing

$$D = \begin{bmatrix} 0 & \omega \\ -\omega & 0 \end{bmatrix} \qquad (4.49)$$

The resulting control policy (4.47) (where $x^3\gamma$, is replaced by the estimates $\hat{x}^3, \hat{\gamma}$ from the state estimator to be designed in a later section), can lend to the closed loop system an insensitivity to the modal damping (and even occasional changes in the damping) that is not enjoyed by the closed loop system in which the mode together with its assumed damping) is <u>not</u> truncated from but included in the model $\mathscr{A}_3$. If the frequency of vibration and the damping are both uncertain then it might be more advantageous to use the error systems of Section 5.0. This is possible because the amplitude of the assumed oscillation is a function of the initial conditions of the synthetic variables (solving (4.43), (4.44) with (4.48))

$$\hat{e}^{31} = C_1(\gamma_o) \sin wt + C_2(\gamma_o) \cos wt$$

Thus, the use of state estimators to update γ_o allows an effective "tracking" of the actual amplitude of oscillation, assuming, in this example, that the damping is small so that the amplitudes change slowly, compared to the τ which we must specify in Section 4.4.

Now that some aspects of my approach to control by model error estimation have emerged, let us pause to clearly state the questions that are before us and see what tasks lie ahead.

1. We wish to model the error system by

$$\hat{e} = P(\tau, d) \gamma$$

$$\dot{\gamma} = D(\tau, d) \gamma$$

Where do we get P, τ, d, and D?

2. We wish to solve the "control" problem

$$u^o = G x^2 \qquad x^2 \triangleq \begin{pmatrix} x^3 \\ \gamma \end{pmatrix}$$

Where do we get G?

3. We wish to design a state estimator for estimating x^2 for feedback. The estimator will have the form

$$\dot{\hat{z}}^t = d_{22} \hat{z}^t + d_{21} z^o + B u^o$$

$$\hat{x}^2 = T_1 \hat{z}^t + T_2 z^o$$

Where do we get d_{22}, d_{21}, B, T_1, T_2?

Now, the specification of G follows in Section 4.3. A way to determine τ is provided in Section 4.4. The structure of the estimator is determined in Section 4.5 and selection of the parameters $(d_{22}, d_{21}, B, T_1, T_2)$ is discussed in Sections 4.6 and 4.7. Alternate choices of D are discussed in Section 5.0, together with special cases of the P selection method which was discussed in Section 4.2.2. Finally d, the order of the error system, is discussed in Section 5.2.

4.3 The Control Problem (Specifying the Matrix G)

In the context of this research exactly what is meant by "the control problem" deserves some clarification. In Figure 2-1, the "control problem" was described as the specification of the matrix G, and that is what we intend to do here. However, we must keep in mind

that the total "controller" design (see Figure 2-1) is not complete until the state estimator is also constructed. Furthermore, both the control law design and the estimator design rely on a model, $\mathcal{A}_2$, as yet incomplete (the "observation window", τ, and the order of the error system, d, which influence the matrices (see (4.41)) $P(\tau,d)$, $D(\tau,d)$ have not been determined). We will not, then, complete our specification of the matrix, G, in this section but we can write the formula for $G(P(\tau,d), D(\tau,d))$.

From the approximation of the control problem (4.32) (4.33) which is given in (4.46), we can write immediately from (4.47)

$$u^o = G \begin{pmatrix} \hat{x}^3 \\ \hat{\gamma} \end{pmatrix} \tag{4.50}$$

where

$$G = - R^{-1} B^{2^T} K^2 \tag{4.51}$$

The Riccati equation (4.48) reduces to, upon substitutions from the definitions under (4.46b),

$$\dot{K} = -KA^3 - A^{3^T}K + KB^3R^{-T}B^{3^T}K - C^{3^T}QC^3 \tag{4.52}$$

$$K(T) = C^{3^T} H\, C^3$$

$$\dot{L} = \left[KB^3R^{-1}B^{3^T} - A^{3^T}\right] L - LD(\tau,d) - K\, P(\tau,d) \tag{4.53}$$

$$L(T) = 0$$

$$\dot{F} = -D^T(\tau,d)F - F\, D(\tau,d) - P^T(\tau,d)L - L^T P(\tau,d) \tag{4.54}$$

$$F(T) = 0$$

where the $n_2 \times n_2$ matrix K^2 has been partitioned according to

$$K^2 = \begin{bmatrix} K & L \\ L^T & F \end{bmatrix} \begin{matrix} \} n_3 \\ \\ \} d \end{matrix} \qquad (4.55)$$

$$\underbrace{}_{n_3} \quad \underbrace{}_{d}$$

It is interesting to note the similarity of these equations (4.52)–(4.54) with those obtained from the singular perturbation approach,[8] where their "model $\mathcal{A}_2$" did not contain synthetic modes but the modes of the untruncated ($\mathcal{A}_1$) model.[*] Note also that (4.52) describes the Riccati matrix for the "error free" $\mathcal{A}_3$ model and can be solved independently of (4.53), (4.54). Given K, then (4.53) is linear in L, and moreover, given L and K (4.54) is linear in F. Finally, from (4.51), together with the definition of B^2 in (4.46), using (4.55),

$$u^o = \left(-R^{-1}B3^T Kx^3 \right) + \left(-R^{-1}B3^T L\gamma \right)$$

$$= u^o(\mathcal{A}_3) + u^o(\hat{e}) \qquad (4.56)$$

Hence, the control consists of the sum of two parts, a term which is optimal for the error free systems, $u^o(\mathcal{A}_3)$, and a term which feeds back a measure of the model error variables, $\gamma(t)$ (recall $\hat{e} = P\gamma$). Now the gain matrix G in (4.50) is given by

$$G = -R^{-1}B3^T [K,L] \qquad (4.57)$$

---

[*]Note that any model we write for a physical system can be considered a truncation from a better model. The point to be made above is that if one considers an $\mathcal{A}_1$ model only slightly larger than the (truncated) $\mathcal{A}_3$ model so that $\mathcal{A}_1$ is of the same order as the $\mathcal{A}_2$ obtained above (by augmenting an assumed small order "error system" to the (truncated) model $\mathcal{A}_3$), then the computational burdens of the two approaches (singular perturbation and model error estimation augmentation) are the same.

There is some question of existence of G when $T \rightarrow \infty$. Since the pairs (A^3, B^3), (A^3, C^3) are respectively, controllable and observable, the existence of a positive definite, symmetric K which satisfies (4.52) for $\dot{K} = 0$ is guaranteed (see for instance, Anderson and Moore,[52] for a discussion of Riccati equations). Therefore, we have only to show that L exists. Consider the defining equation for L, (4.53),

$$\dot{L} = -LD - [A^3 + B^3 G_1] L - KP \qquad L(T) = 0 \qquad (4.58)$$

where $G_1 \triangleq - R^{-1} B^{3T} K$ is the gain for the error free system. Now we show that L(0) exists if the error free system

$$\dot{x}^3 = [A^3 + B^3 G_1] x^3$$

is sufficiently stable. To show this, note that if L(t) has a steady state solution, L(0), which satisfies

$$0 = L(0)D - [A^3 + B^3 G_1] L(0) - KP \qquad (4.59)$$

then the difference

$$\psi(t,T) \triangleq L(t,T) - L(0) \qquad (4.60)$$

satisfies (differentiate (4.60), using (4.59), (4.58))

$$\dot{\psi}(t,T) = -(A^3 + B^3 G_1) \psi(t,T) - \psi(t,T)D \qquad (4.61)$$

where the notation $L(t,t_1)$ represents the solution of a linear matrix equation such as (4.58), expressed by the matrix variation of constants formular (see Brockett[40] p. 59)

$$L(t,t_1) = \Phi^1(t,t_1)L(t_1) \Phi^2(t,t_1) - \int_t^{t_1} \Phi^1(t,\sigma) K(\sigma)P\Phi^2(t,\sigma)d\sigma$$

$$(4.62)$$

where Φ^1 is the state transition matrix for $[-A^3 - B^3 G_1]$ and Φ^2 is the

66

state transition matrix for [-D]. Now, again applying the matrix variation of constants formula to (4.61) we have

$$\psi(t,T) = \Phi^1(t,T) \ [L(T,T) - L(0)]\Phi^2(t,T) \tag{4.63}$$

or simply, for constant $[A^3 + B^3G]$, D,

$$L(t,T) = L(0) - e^{[A^3+B^3G_1](T-t)} L(0) \ e^{D^T(T-t)}$$

Thus $L(0)$ exists only if $\displaystyle\lim_{T \to \infty} L(t,T)$ exists which is guaranteed if

$$\text{Re } \lambda_i[D] + \text{Re } \lambda_j[A^3 + B^3G_1] < 0 \tag{4.64}$$

$\forall i \in [1,d]$, $j \in [1,n_3]$. Thus, the existence of L is guaranteed under condition (4.64). Now we turn to the task of defining the "observation window" τ, used in (4.41), and the overall design of the state estimator for $\hat{x}^3$, $\hat{\gamma}$.

4.4 Specifying the Observation Window, τ.

In the operation of a state estimator (Kalman or Luenberger) three basic functions are performed: 1) the receiving of measurements, 2) the prediction* of measurements, and 3) the feedback of the difference between the measurements and their predictions. The first function causally relates the estimates to the physical system. The second function relies upon the model of the system, built within the estimator, to instantaneously predict measurements, so that the estimate of

---

*The use of the word "prediction" is not here restricted to the stochastic estimation theory. For instance, in the deterministic estimator, the quantity $\hat{z}^j = M^j \ \hat{x}^j$, where $\hat{x}^j$ is the output of the estimator, is called a prediction (instantaneous) of the measurement (based upon model $\mathcal{L}_j$).

the state vector can be changed (the third estimator function) if there
is a difference between this prediction, $\hat{z}^j = M^j\hat{x}^j$, and the real
measurement, $z^o(t)$. (The difference $z^o - z^j$ is called the measurement
residual, $\tilde{z}^j$, see (4.29).) In other words, unless there is a change in
the "initial conditions" of the physical system the state estimator will
not change its estimate of the state <u>assuming</u> the model within the esti-
mator is correct (so that the measurement prediction, $M^j\hat{x}^j(t)$, remains
equal to the measurement, $z^o(t)$). Now if one is concerned about con-
structing a model upon which to base the design of the estimator, a
logical question to ask is "Over what interval of time, τ, should the
model (predictions, $z^j = M^j x^j(t)$) be accurate?". That is, how long
should the zero state response of the model match (to some prespecified
accuracy) the zero state response physical system (see Figure 4-1 for
the "experiment")? Let us first agree on two extremes. If this time,
τ, is "very long" then in some sense the model is "very good". In
fact, if one performed some (off-line) performance analysis with this
model, to determine some appropriate controls, one might, on the basis
of the above experiment, decide to issue such controls to the physical
system in an open loop manner (without further use of the model in an
on-line operation, having previously verified by the experiment that
the model predictions are "very good"). Furthermore, if one used this
model to implement a state estimator for on-line feedback control, then
under the same conditions of the experiment the measurement residual
would (after an initial transient of the stable estimator) remain near
zero, $z^o \approx M^j\hat{x}^j$, and the feedback signal would soon be based solely on

the prediction of the model.[*] On the other extreme, if the time, τ,
in the above experiment is very short and approaches zero then the
model is arbitrarily "bad" in the sense that no model can be worse,
since there is zero "correlation time" with the physical system.
Intuitively, one would guess that for stability of the closed loop
system the estimator (model) need only be accurate over time constants
of the closed loop system. This discussion prompts two definitions
which might be used as a characteristic property of the model. We
define a property called a "characteristic time".

Definition 4.1: Given a linearized time invariant model, $\mathscr{A}_j$ of a
physical system, $\mathscr{A}_o$,

$$\mathscr{A}_j \begin{cases} \dot{x}^j = A^j\,x^j + B^j\,u^o \\ z^i = M^j\,x^j \end{cases}$$

Define a characteristic time, τ^j, associated with the model $\mathscr{A}_j$ such that
the zero state response of the model, $z^j_{ZSR}(t)$, satisfies

$$||z^o(t) - z^j_{ZSR}(t)|| \leq \delta \quad \forall\, \delta \in [0,\tau^j]$$

if $\qquad ||z^o(0) - z^j_{ZSR}(0)|| = ||z^o(0)|| \leq \delta_o,$[**]

see Figure 4-1. This definition of characteristic time of a model is

_____

[*] Alternately, it can happen in Kalman filters, with small or zero
assumed plant noise (see Section 3), that the weighting (Kalman gain)
on the measurement residual goes to zero because the estimator "thinks"
it has learned the model and disconnects itself ($\hat{K} \to 0$) from the
(noisy) measurements, basing all future state estimates solely on the
prediction of the model. The result, of course, is filter divergence,
since the model is never perfect.

[**] To make this definition practical in a deterministic sense some toler-
ance must be accepted on the initial condition $z^o(0)$ of the physical
system, since this cannot necessarily be controlled to absolute zero.
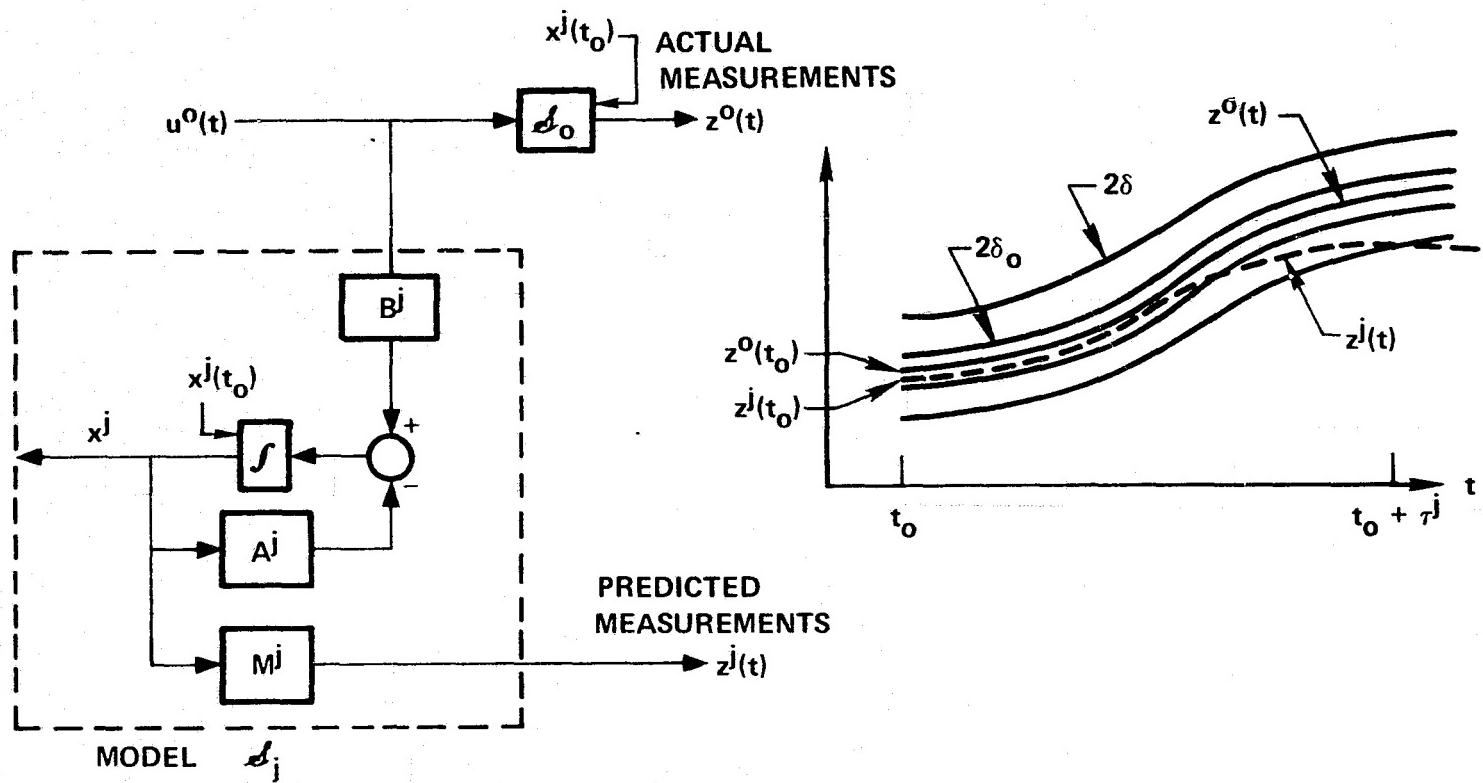
Figure 4-1. Characteristic Times of Models.

not very useful owing to its empirical nature and the fact that it is a function of the small parameters δ_o, δ. It is, however, closer to the true property of the model (relative to the physical system) that is important to us in estimator design than is the more tractable definition which follows. Note that the above τ is not a property solely of the model. That is, τ^j cannot be determined solely by examination of the model parameters (such as eigenvalues of the matrix A^j). The following is a definition of τ which is determined solely from parameters of the model and the control law. It will be taken as the definition of model characteristic time throughout the remainder of the report.

Definition 4.2: Let $V(x^j(t))$ be a Liapunov function for the asymptotically stable model of the closed loop system

$$\mathscr{A}_j' \quad \begin{cases} \dot{x}^j = A^j x^j + B^j u^o(x^j), & u^o(x^j) \text{ given} \\ y^j = C^j x^j \end{cases} \tag{4.65}$$

Then the quantity

$$\tau^j \triangleq \quad \max_{x^j} \left[\frac{V(x^j(t))}{-\dot{V}(x^j(t))} \right] \tag{4.66}$$

will be called a characteristic time of the model $\mathscr{A}_j$.

This definition of a characteristic time was first suggested by Kalman, Reference 53. The proof that there exists such a τ^j rests on the proof of the fact that there exists, for any system $\mathscr{A}_j'$ which is asymptotically stable, a scalar function $v(x^j(t))$ which is positive definite and for which $\dot{v}(x^j(t))$ is negative definite. Such a proof was provided by Messera, see Theorem 23 of Reference 54. It is another matter,

however, to find such a function. In the following discussion let us suppose there is such a Liapunov function of quadratic form

$$V(x^j) = x^{j^T} K^j x^j \tag{4.67}$$

for the closed loop system model $\mathcal{A}'_j$ which is linear.

$$\mathcal{A}'_j \quad \left\{ \begin{array}{l} \dot{x}^j = \bar{A}^j x^j \\ y^j = C^j x^j \end{array} \right\} \tag{4.68}$$

where

$$\bar{A}^j \triangleq A^j + B^j G^j$$

$$G^j \triangleq \text{state feedback gain matrix.}$$

We have tacitly assumed here that the state variables are available for feedback. This is tantamount to the assumption that the state estimator is perfect, an assumption which we will not tolerate elsewhere but in this definition of τ. Now we cite the useful

Theorem 4.1: (Using some results in Reference 40 p. 128). Let the closed loop system (4.65) have the Liapunov function (4.67) which satisfies the Liapunov theorem for asymptotic stability (Theorem 1 of Reference 53). Then $\dot{V}(x^j)$ is a negative definite function of x^j and the model characteristic time, τ, defined by (4.66), is given by the spectral radius of the matrix $[R^{j^{-1}} K^j]$, where $R^j \triangleq \bar{A}^{j^T} K^j + K^j \bar{A}^j$.

Proof: Extreme values of the scalar function

$$\tau(x^j) = \frac{V(x^j(t))}{-\dot{V}(x^j(t))} \tag{4.69}$$

must satisfy

$$\frac{\partial \tau(x^j)}{\partial x^j} = 0 \tag{4.70}$$

From (4.67), (4.68), (4.69),

$$\frac{\partial \tau(x^j)}{\partial x^j} = \frac{x^{j^T}(K^j + K^{j^T})}{x^{j^T} R^j x^j} - \frac{x^{j^T}(R^j + R^{j^T})}{(x^{j^T} R^j x^j)^2} x^{j^T} K^j x^j = 0 \tag{4.71}$$

multiplying by $\frac{1}{2} x^{j^T} R^j x^j$, we have, for asymmetric K^j,

$$x^{j^T}[K^j - \tau(x^j)R^j] = 0 \tag{4.72}$$

which implies that $\tau(x^j)$ must be an eigenvalue of the matrix $R^{j^{-1}} K^j$.
Hence

$$\tau^j \triangleq \max_{x^j} \tau(x^j) = \lambda_{max}[R^{j^{-1}} K^j]$$
$$\triangleq r[R^{j^{-1}} K^j] \tag{4.73}$$

where $\lambda_{max}[\cdot]$ denotes maximum eigenvalue (magnitude) of $[\cdot]$, sometimes called the spectral radius of $[\cdot]$, denoted $r[\cdot]$.

Corollary: If $V(x^j)$ is chosen for the system (4.65), as

$$V(x^j) \triangleq \int_t^\infty (y^{j^T} Q y^j + u^{o^T} R u^o) \, dt \tag{4.74}$$

and $u^o(x^j)$ is selected as

$$u^o(x^j) = -R^{-1} B^{j^T} K^j x^j \tag{4.75}$$

where K^j is the symmetric positive definite solution of

$$0 = -K^j A^j - A^{j^T} K^j + K^j B^j R^{-1} B^{j^T} K^j - C^{j^T} Q C^j \tag{4.76}$$

and where Q and R are positive definite matrices and the pairs (A^j, B^j), (A^j, C^j) are respectively controllable and observable, then the matrix $\bar{A}^j$ in (4.68) becomes

$$\bar{A}^j = A^j - B^j R^{-1} B^{j^T} K^j \tag{4.77}$$

and (4.73) becomes, using (4.77), (4.76),

$$\tau^j = \lambda_{max} \left\{ -\left[K^j B^j R^{-1} B^{j^T} K^j + C^{j^T} Q C^j \right]^{-1} K^j \right\} \quad (4.78)$$

where we have used the fact that (4.74) may be expressed $V(x^j) =$
$x^{j^T} K^j x^j$; see p. 277 of Reference 16 for proof of this equivalent
expression for $V(x^j)$. The minus sign in (4.78) can be neglected since
λ_{max} was defined to mean the eigenvalue with maximum magnitude,
$\lambda_{max} = |\sigma + j\omega|_{max}$. The "observation window", τ, needed for the model
$\mathcal{A}_2$ in (4.46b) for the estimator design is selected as

$$\tau \triangleq \tau^3$$

where τ^3 is defined by (4.78) with $j = 3$. In this way the time con-
stants of the truncated model, $\mathcal{A}_3$, are used to influence the error
systems (4.43), (4.44) which is augmented to $\mathcal{A}_3$ to define $\mathcal{A}_2$ ($\mathcal{A}_2$ is to
be used for the design of the structure (order) of the controller
(estimator and control law)).

4.5 Design of the State Estimator and Special Cases

The problem to be solved in this section is the problem imposed by
the decision to implement a control policy which requires all n_j state
variables of the model $\mathcal{A}_j$ (such as $u^o = Gx^j$, $x^j \in R^{n_j}$) when there are
only ℓ measurements available, $z^o(t)$, $\dot{z} \in R^\ell$. The relationship of the
resulting design of a minimal order $(n_j - \ell)$ state estimator with the
standard Luenberger and Kalman estimators is discussed. In particular,
the unifying approach of carrying along "model error vectors" in the
design allows one to later consider the errors to be either determin-
istic or random (white noise) and to obtain, respectively, a Luenberger

observer for the augmented system $\mathcal{A}_2$ ($\mathcal{A}_3$ plus error system), or a Kalman filter in the limiting case of vanishing measurement noise. When the $\mathcal{A}_3$ model is assumed perfectly accurate ($\mathcal{A}_0 = \mathcal{A}_3$), then the standard Luenberger observer is obtained. In this way the control problem need not be considered either as "deterministic" or "stochastic" until the selection of estimator gains is made, and moreover, the choice may be made under the consideration of both deterministic and stochastic characterizations of the error vectors.

In section 4.1 we discussed the existence of a linear coordinate transformation which would transform the large state vector x^1 into a vector of variables ξ in which x^3 appeared as a partitioned subvector of ξ where x^3 were certain variables whose physical significance was known to us. Here we use the same transformation device.

We have available to us certain measurements $z^0(t)$, which in model $\mathcal{A}_j$ have been described approximately by $z^j = M^j x^j$.

We wish to obtain an approximation of the vector x^j which has a larger dimension than z^0. Consider therefore a nonsingular coordinate transformation

$$x^j = T \xi \tag{4.79}$$

such that if x^j obeys

$$
\begin{aligned}
\dot{x}^j &= A^j x^j + B^j u^0 + W & x^j &\in R^{n_j} \\
z^j &= M^j x^j & z^j &\in R^{\ell} \\
y^j &= C^j x^j & u^0 &\in R^m
\end{aligned}
\tag{4.80}
$$

then ξ obeys

75

$$\dot{\mathcal{X}} = T^{-1} A^j T \mathcal{X} + T^{-1} B^j u^o + T^{-1} w$$

$$z^j = M^j T \mathcal{X} \qquad\qquad (4.81)$$

$$y^j = C^j T \mathcal{X}$$

Furthermore, let T be that special transformation which reveals the measurement variables, z^j, as explicit state variables (say the first ℓ' of them) of the transformed system (4.81). Then

$$\mathcal{X} = \begin{pmatrix} z^j \\ z^t \end{pmatrix} \qquad \begin{matrix} z^j \in R^{\ell'}, \\ z^t \in R^{n_j - \ell} \end{matrix} \qquad 0 \le \ell' \le \ell \qquad (4.82)$$

where z^t is an $(n_j - \ell')$ vector of "truncated measurements" in the sense that they are not directly measurable from the system (4.80), as are z_i^j. However, since we have available (given T) the differential equation (4.81) which z^t obeys we suspect that under certain conditions we can, by integrating the differential equation in z^t, obtain an approximation of $z^t(t)$. In fact, since we have the differential equation for z^t, the only question is whether we can recover the initial conditions, $z^t(0)$. Now in the transformed coordinates of (4.81) we have

$$\left.\begin{aligned}
\dot{z}^j &= d_{11} z^j + d_{12} z^t + B_1 u^o + {}_1 w \\
\dot{z}^t &= d_{21} z^j + d_{22} z^t + B_2 u^o + {}_1 w \\
y^j &= c_1 z^j + c_2 z^t \\
z^j &= M_1 z^j + M_2 z^t
\end{aligned}\right\} \qquad (4.83)$$

where

$$\begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix} \triangleq T^{-1} A^j T$$

$$\begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = T^{-1} B^j \quad , \quad \begin{bmatrix} E_1 \\ E_2 \end{bmatrix} = T^{-1}$$

$$[C_1, C_2] = C^j T$$

$$[M_1, M_2] = M^j T$$

Partition the matrix T so as to define the matrices T_{ij}, Γ_{ij} as follows.

$$T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{matrix} \} \ell' \\ \} n_j - \ell' = q \end{matrix}$$

$$\underbrace{\qquad}_{\ell'} \quad \underbrace{\qquad}_{q}$$

$$T^{-1} = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{bmatrix} \begin{matrix} \} \ell' \\ \} q \end{matrix}$$

(4.84)

Recall that eigenvalues are preserved through a similarity transformation. Thus the eigenvalues of $T^{-1} A^j T$ are fixed a priori by the eigenvalues of A^j. Now the fundamental principle upon which reduced order observers (state estimators) rely is that even though the eigenvalues of $T^{-1} A^j T$ are fixed a priori, it is possible under certain conditions to select T such that the eigenvalues of a block diagonal submatrix within $T^{-1} A^j T$ may be arbitrarily specified. The conditions under which this is true and the relevance of this fact to the design of the state estimators are developed in what follows. To simplify our immediate task we will only consider the case $\ell' = \ell$. It turns out that $\ell' = \ell$ yields a minimal order state estimator. Circumstances

when one might not wish to design the minimal order estimator are discussed in the sequel.

Two requirements on the transformation T are immediate for non-singularity. From (4.84),

$$T^{-1} T = I \implies \begin{cases} \Gamma_{11}\, T_{11} + \Gamma_{12}\, T_{21} = I_{\ell'} \\[1em] \Gamma_{11}\, T_{12} + \Gamma_{12}\, T_{22} = 0_{\ell' q} \\[1em] \Gamma_{21}\, T_{11} + \Gamma_{22}\, T_{21} = 0_{q\ell'} \\[1em] \Gamma_{21}\, T_{12} + \Gamma_{22}\, T_{22} = I_q \end{cases} \qquad (4.85)$$

$$TT^{-1} = I \implies \begin{cases} T_{11}\, \Gamma_{11} + T_{12}\, \Gamma_{21} = I_{\ell'} \\[1em] T_{11}\, \Gamma_{12} + T_{12}\, \Gamma_{22} = 0_{\ell' q} \\[1em] T_{21}\, \Gamma_{11} + T_{22}\, \Gamma_{21} = 0_{q\ell'} \\[1em] T_{21}\, \Gamma_{12} + T_{22}\, \Gamma_{22} = I_q \end{cases} \qquad (4.86)$$

and from (4.82), (4.79), and (4.84)[*]

$$\left[\Gamma_{11} , \Gamma_{12} \right]_{\ell n_j} = M^j \qquad (4.87)$$

For convenience define, T_1, T_2, Γ by

$$T_1 \triangleq \begin{bmatrix} T_{11} \\ T_{21} \end{bmatrix}_{n_j \ell} \qquad T_2 = \begin{bmatrix} T_{12} \\ T_{22} \end{bmatrix}_{n_j q}$$

$$\Gamma = \begin{bmatrix} \Gamma_{21} , \Gamma_{22} \end{bmatrix}_{q n_j}$$

Then

―――――――――

[*]The notation $[\cdot]_{jk}$ means that the matrix $[\cdot]$ has dimensions r by k.

$$T^{-1} A^j T = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix} = \begin{bmatrix} M^j \\ \Gamma \end{bmatrix} A^j \begin{bmatrix} T_1 & T_2 \end{bmatrix}$$

$$= \begin{bmatrix} M^j A^j T_1 & M^j A^j T_2 \\ \Gamma A^j T_1 & \Gamma A^j T_2 \end{bmatrix} \tag{4.88}$$

Also (4.85), (4.86) can be expressed in the form

$$T^{-1} T = \begin{bmatrix} M^j T_1 & M^j T_2 \\ \Gamma T_1 & \Gamma T_2 \end{bmatrix} = \begin{bmatrix} I_\ell & 0 \\ 0 & I_q \end{bmatrix} \tag{4.89a}$$

$$T T^{-1} = T_1 M^j + T_2 \Gamma = I_{n_j} \tag{4.89b}$$

Let us consider (4.79) as the output relationship for the transformed system

$$\left. \begin{aligned} \dot{\text{x}} &= T^{-1} A^j T \, \text{x} + T^{-1} B^j u^o \\ x^j &= T \, \text{x} = T_1 z^j + T_2 z^t \end{aligned} \right\} \tag{4.90}$$

Now if we integrate (4.90) for any given $u^o(t)$ we have $x^j(t)$ exactly, assuming $\text{x}(0) = T^{-1} x^1(0)$, because (4.90) is just a transformation of the defining relationship for $x^j(\text{x})$, namely (4.80). On the other hand, suppose we integrate only the z^t equation in (4.90), shown more specifically in (4.83). Now, instead of taking z^j (which is needed to solve for $z^t(t)$ in (4.83)) as the integral of the first equation in (4.83), suppose we take instead the measurements available from the physical system, $z^o(t)$, and construct (define) the estimate $\hat{x}^j$ as follows

$$\dot{\hat{z}}^t = d_{22}\, \hat{z}^t + d_{21}\, z^o + B_2\, u^o$$

$$\hat{x}^j = T_1\, z^o + T_2\, \hat{z}^t \qquad\qquad (4.91)$$

In order to see how closely $\hat{x}^j(t)$ approximates $x^j(t)$, which is defined by (4.80) (or equivalently and more conveniently by (4.90) with $\chi(0) = T^{-1} x^j(0)$), let us rewrite (4.90) relating z^j (that measurement which is predicted on the basis of the model $\mathcal{A}_j$) to the actual measurement $z^o(t)$, as follows

$$z^j = z^o - \tilde{z}^j \qquad \left(\tilde{z}^j \triangleq z^o - z^j \right)$$

where $\tilde{z}$ is a measurement residual and is present because of errors in the model $\mathcal{A}_j$. Then (4.90) becomes

$$\left.\begin{aligned}
\dot{z}^o - \dot{\tilde{z}}^j &= d_{11}(z^o - \tilde{z}^j) + d_{12}\, z^t + B_1\, u^o \\
\dot{z}^t &= d_{21}(z^o - \tilde{z}^j) + d_{22}\, z^t + B_2\, u^o
\end{aligned}\right\} \qquad (4.92)$$

$$x^j = T_1(z^o - \tilde{z}^j) + T_2\, z^t$$

Defining the estimator error vector

$$\tilde{z}^t \triangleq z^t - \hat{z}^t \qquad\qquad (4.93)$$

and the estimation error vector

$$\tilde{x}^j \triangleq x^j - \hat{x}^j \qquad\qquad (4.94)$$

it is straightforward to show, by differentiating (4.93), using (4.91), (4.92) and (4.88), that $\tilde{z}^t(t)$ obeys

$$\dot{\tilde{z}}^t = d_{22}\, \tilde{z}^t - d_{21}\, \tilde{z}^j = \Gamma A^j T_2 \tilde{z}^t - \Gamma A^j T_1\, \tilde{z}^j \qquad (4.95)$$

and likewise for $\tilde{x}^j$

$$\dot{\tilde{x}}^j = T_2 \, \Gamma \, A^j \, \tilde{x}^j - T_1 \, \dot{\tilde{z}} \qquad\qquad (4.96)$$

Note also that

$$\tilde{z}^t = \Gamma \, \tilde{x}^j \qquad\qquad (4.97)$$

The requirements for perfect estimation and asymptotic estimation are now before us. If we view (4.97) as the output of the system (4.96) we can see that if the pair $(T_2 \Gamma A^j, \Gamma)$ is observable, then the event $\tilde{z}^t(t) \equiv 0$ guarantees that $\tilde{x}^j(t) \equiv 0$. Therefore, we conclude that the estimate of $x^j(t)$ generated by (4.91) can be made to approach $x^j(t)$ arbitrarily fast if a transformation matrix T can be found such that

1. T is nonsingular ((4.89) is satisfied)

2. The eigenvalues of the submatrix $d_{22} \triangleq [T^{-1} A^j T]_{22}$ may be arbitrarily specified

3. The measurements appear as state variables in the transformed equations (that is, (4.109) is satisfied)

and the final requirement is that the model $\mathscr{A}_j$ is perfect in the sense that $\tilde{z}^j \triangleq z^0 - z^j \equiv 0$. These conditions are further summarized in the

Definition: The $(n_j - \ell)^{\text{th}}$ order system of (4.91) is said to be a state estimator for the model (4.80) if and only if

1) $\tilde{z}^j \equiv 0$ (the model is perfect)

2) Γ is selected to satisfy (4.89) with the additional requirement that $d_{22} \triangleq \Gamma A^j T_2$ is a stability matrix.

To show that these requirements for an estimator are the same as those of Luenberger, multiply the equation in (4.89b) by ΓA^j and compare with Equation (5.5a) in Reference 55. Note also that the requirement 4), p. 275 of Kwatny,[56] is equivalent to (4.89b) but the additional

requirement, 2), of Kwatny is unnecessary since 2) can be obtained by multiplying 4) by ΓA^j. Now following a suggestion of Reference 57, p. 606, to make Γ_{11} in (4.87) nonsingular, we shall eliminate some of the unknowns in (4.85) to facilitate design of the estimator. Suppose the state variables of (4.80) are arranged so that Γ_{11} in (4.87) is nonsingular. Then from (4.85) it follows that T_{11} and T_{12} can be eliminated, yielding

$$T = \begin{bmatrix} \Gamma_{11}^{-1}(I - \Gamma_{12}T_{21}) & -\Gamma_{11}^{-1}\Gamma_{12}T_{22} \\ T_{21} & T_{22} \end{bmatrix} \tag{4.98}$$

which is nonsingular only if T_{22} is nonsingular, and thus

$$T^{-1} = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ -T_{22}^{-1}T_{21}\Gamma_{11} & T_{22}^{-1}\left[I - T_{21}\Gamma_{12}\right] \end{bmatrix} \tag{4.99}$$

From (4.88)

$$T^{-1}A^j T = \begin{bmatrix} \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \end{bmatrix}A^j \begin{pmatrix} \Gamma_{11}^{-1}\left[I-\Gamma_{12}T_{21}\right] \\ T_{21} \end{pmatrix} & \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \end{bmatrix}A^j \begin{bmatrix} -\Gamma_{11}^{-1}\Gamma_{12} \\ I \end{bmatrix}T_{22} \\ \begin{bmatrix} \Gamma_{21} & \Gamma_{22} \end{bmatrix}A^j \begin{pmatrix} \Gamma_{11}^{-1}\left[I-\Gamma_{12}T_{21}\right] \\ T_{21} \end{pmatrix} & \begin{bmatrix} \Gamma_{21} & \Gamma_{22} \end{bmatrix}A^j \begin{bmatrix} -\Gamma_{11}^{-1}\Gamma_{12} \\ I \end{bmatrix}T_{22} \end{bmatrix}$$

$$\tag{4.100}$$

where,

$$\Gamma_{21} = -T_{22}^{-1}T_{21}\Gamma_{11}$$

$$\Gamma_{22} = T_{22}^{-1}\left[I - T_{21}\Gamma_{12}\right]$$

After some manipulation it is possible to write the neater form

$$T^{-1}A^jT = A = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix} = \left[\begin{array}{c|c} M^jA^jL_1 & M^jA^jL_2\,T_{22} \\ \hline T_{22}^{-1}\left[F_1 + T_{21}\,G_1\right] & T_{22}^{-1}\left[F_2 + T_{21}\,G_2\right]T_{22} \end{array} \right]$$

(4.101)

where

$$L_1 \triangleq \begin{bmatrix} \Gamma_{11}^{-1}\left[I - \Gamma_{12}\,T_{21}\right] \\ T_{21} \end{bmatrix} \begin{matrix} \ell \\ n_j - \ell = q \end{matrix} \quad = T_1 \qquad (4.102)$$

$$L_2 \triangleq \begin{bmatrix} -\Gamma_{11}^{-1}\,\Gamma_{12} \\ I_q \end{bmatrix} \begin{matrix} \ell \\ q \end{matrix} \quad = T_2\,T_{22}^{-1} \qquad (4.103)$$

$$F_1 \triangleq \begin{bmatrix} 0_{q\ell} & , & I_q \end{bmatrix}A^j\,L_1 \qquad (4.104a)$$

$$F_2 \triangleq \begin{bmatrix} 0_{q\ell} & , & I_q \end{bmatrix}A^j\,L_2 \qquad (4.104b)$$

$$G_1 \triangleq -M^j\,A^j\,L_1 \qquad (4.104c)$$

$$G_2 \triangleq -M^j\,A^j\,L_2 \qquad (4.104d)$$

$$\begin{bmatrix} \Gamma_{11} & , & \Gamma_{12} \end{bmatrix} \triangleq M^j \qquad (4.104e)$$

Now, T_{22} serves only as a transformation of estimator coordinates, as seen by noting that the eigenvalues of d_{22} are the eigenvalues of $[F_2 + T_{21}\,G_2]$. Since the eigenvalues of d_{22} (and of the estimator; see (4.91)) are independent of the required nonsingular matrix T_{22}, we may without loss of generality set, hereafter,

83

$$T_{22} = I$$

and write

$$\begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix} = \left[\begin{array}{c|c} M^j A^j L_1 & M^j A^j L_2 \\ \hline F_1 + T_{21} G_1 & F_2 + T_{21} G_2 \end{array} \right] \qquad (4.105)$$

By a fundamental theorem of linear system theory, the matrix $F_2 + T_{21} G_2$ has an arbitrary spectrum if and only if the pair (F_2, G_2) is observable, or equivalently the pair (F_2^T, G_2^T) is controllable (see Wonham[58], p. 660 for a proof). Also, the structure of F_2 and G_2 is such that the pair (F_2, G_2) is observable only if the pair (A^j, M^j) is observable. (This can be shown by construction using the definition of observability).

4.6 Estimator Gain Selection on the Basis of Model $\mathscr{A}_2$

One approach for selecting the free parameters, T_{21} in the design of the estimator, (4.91), is to make the estimator stable with respect to model $\mathscr{A}_2$. In this event the stability properties of the matrix d_{22} in (4.105) are of interest. Another approach is to relate T_{21} to the properties of the closed loop system using model $\mathscr{A}_1$. This latter approach will be discussed in section 4.7. Presently, we proceed to select T_{21} with reference to model $\mathscr{A}_2$ (that is, select T_{21} in (4.105) to make d_{22} stable).

One approach for selecting the gains of a state estimator proposed in the Appendix of Johnson,[3] requires no transformation to canonical form as do the algorithms proposed in References 57, 59, and 60. The approach also leads to a state estimator which is equivalent to a Kalman filter in the limiting case of vanishing measurement noise.

84

In other words, the following completely deterministic design is also optimal for the linear stochastic problem driven by white Gaussian disturbances for some covariance of the plant. Furthermore, the relationships between the covariance of the plant noise and the parameters of the deterministic estimator design are given explicitly in (21) of Reference 56. One can therefore proceed to design any minimal order $(n_j - \ell)$ state estimator on the basis of a given covariance matrix of the plant noise or, on the basis of specified stability properties of the estimator, whichever information/criterion seems more credible in a particular application. It also readily follows that any linear state estimator of order n_j is equivalent to a Kalman filter in the steady state for some plant white noise covariance matrix and some measurement noise covariance matrix (see Reference 61). There are a number of the reduced order state estimator designs for stochastic problems by Novak,[62] such as those proposed, which allow the order of the estimator to be specified anywhere between n_j and $n_j - \ell$.

In general, the following conjecture seems appropriate.

Conjecture: Any stable linear state estimator of order n, $n_j \geq n \geq n_j - \ell$, is equivalent to a Kalman estimator for some given covariance matrix of the plant disturbance process and some given measurement covariance matrix of rank $(\ell + n - n_j)$. The two limiting cases of this conjecture $(n = n_j - \ell$, and $n = n_j)$ have been verified in the literature; see Reference 56 and 61, respectively. Note that an $\ell \times \ell$ measurement covariance matrix of rank $r < \ell$ means that $\ell - r$ measurements (or linear combinations thereof) are deterministic (known exactly). We

have now supporting arguments for a previous assertion of section 4.5, that a linear control problem need not be viewed as either determinstic or stochastic until the selection of the state estimator gains must be accomplished.

A matrix T_{21} which will make d_{22} in (4.105) a stability matrix may be obtained as follows. Consider the optimal control problem

$$\min_{\omega} \int_0^\infty (\rho^T \rho + \omega^T R_\xi \omega) dt$$

subject to

$$\dot{\xi} = F_2^T \xi + G_2^T \omega$$

$$\rho = Q_\xi \xi$$

Then if the pairs (F_2, G_2), (F_2^T, Q_ξ) are both observable,[*] the matrix for the closed loop system

$$\dot{\xi} = \left[F_2^T - G_2^T R_\xi^{-1} G_2 K \right] \xi \qquad (4.106)$$

has an arbitrary spectrum, where K is the symmetric positive definite solution of

$$0 = - K F_2^T - F_2 K + K G_2^T R_\xi^{-1} G_2 K - Q_g^T Q_g \qquad (4.107)$$

Therefore the choice

$$T_{21} = - K G_2^T R_g^{-1} \qquad (4.108)$$

yields a stable estimator (4.91) for any positive definite matrix R_ξ. (Compare the matrix in (4.106) with d_{22} in (4.105) and recall that a matrix and its transpose share the same spectrum). Furthermore, if

---

[*] Recall that observability of the pair (F_2, G_2) is equivalent to controllability of the pair (F_2^T, G_2^T).

the estimator is to have time constants smaller than the "observation window," τ, specified in section 4.4, this can be accomplished by substituting $[F_2 + 1/\tau \; I]$ for F_2 in (4.107).

By using some results from Kwatny[56] it can be shown that the procedure described above (4.107), (4.108) for selecting estimator gains yields an estimator which is also optimal in the stochastic sense with some restrictions: The (deterministic) estimator designed above is also a Kalman filter in the case of vanishing measurement noise. Kwatny shows the explicit relationships between the plant noise covariance matrix assumed for the stochastic problem and the deterministic design parameters such as R_ξ, Q_ξ in (4.107), (4.108). We may therefore fix the estimator gains on the basis of either deterministic considerations (stability of estimator) or stochastic considerations (minimum variance estimator) simply by making an appropriate choice for Q_ξ and R_ξ in (4.107, (4.108).

To model the behavior of the closed loop system assume that the control law is a linear combination of the estimates of the states of model $\mathcal{A}_2$.

$$\text{control} \atop \text{law} \left\{ u^o(t) = G \; \hat{x}^2(t) \right. \tag{4.109}$$

On the basis of the state estimator designed, at least in structure, in section 4.6 we have from (4.91),

$$\text{state} \atop \text{estimator} \left\{ \begin{array}{l} \dot{\hat{z}}^t = d_{22} \; \hat{z}^t + d_{21} \; z^o + B_2 \; G(T_1 \; z^o + T_2 \; \hat{z}^t) \\ \hat{x}^2 = T_1 \; z^o + T_2 \; \hat{z}^t \end{array} \right. \tag{4.110}$$

Now, therefore, the dynamical system which is composed of the control law (4.109) together with the state estimator (4.110) is called the

"controller" and is given by,

controller $\begin{cases} \dot{\hat{z}}^t = [d_{22} + B_2 GT_2]\hat{z}^t + [d_{21} + B_2 GT_1]z^o \\ u^o(t) = GT_1 \ z^o + GT_2 \ \hat{z}^t \end{cases}$ (4.111)

The controller receives as inputs the physical measurements, $z^o)(t)$, and yields as its outputs the controls $u^o(t)$ which are to be applied to the physical system.

To evaluate a candidate controller design one could attach this controller to the physical system which actually receives $u^o(t)$ and generates $z^o(t)$. Such an experiment with the physical system might be economically prohibitive. Suppose for instance, that the gravitational environment on the earth's surface precluded ground test of a fine pointing system of a spacecraft; or that a mass transit control policy evaluation imposes unacceptable risk of life and property. In such cases evaluations are performed on the basis of mathematical models of the physical system rather than the physical system itself. Such a prediction of physical system performance on the basis of a mathematical model which has been programmed on a (digital or analog) computer is called a "simulation." It is really this set of circumstances which provides motivation in this research to introduce "multimodel" designs. By multimodel designs we mean that selection of design parameters and controller structure is not based upon the same model. The actual control law design might be (wisely) based upon one model, call it $\mathcal{A}_2$, whereas the estimator gains might be chosen on the basis of still another (better) model, $\mathcal{A}_1$, which might be used for evaluation of the total controller design. The simpler model $\mathcal{A}_2$ serves to fix the structure of the controller (i.e., order of

88

estimator) whereas the information available in a better model, $\mathscr{A}_1$ serves to make the state estimates (of that simpler model) more accurate.

To see if this is practical let us construct the closed loop system with model $\mathscr{A}_1$ using the controller (4.111), which was designed on the basis of model $\mathscr{A}_2$. Add to the model $\mathscr{A}_1$,

$$\mathscr{A}_1 \begin{cases} \dot{x}^1 = A^1 x^1 + B^1 u^o + w^1 \\ \begin{pmatrix} y^1 \\ z^1 \end{pmatrix} = \begin{bmatrix} C^1 \\ M^1 \end{bmatrix} x^1 \end{cases} \qquad (4.112)$$

the controller (4.111), to obtain

$$\mathscr{A}_1' \begin{cases} \begin{pmatrix} \dot{x}^1 \\ \dot{z}^t \end{pmatrix} = \begin{bmatrix} A^1 & B^1 GT_2 \\ \hline 0 & d_{22} + B_2 GT_2 \end{bmatrix} \begin{pmatrix} x^1 \\ z^t \end{pmatrix} + \begin{bmatrix} B^1 GT_1 \\ d_{21} + B_2 GT_1 \end{bmatrix} z^o + \begin{bmatrix} I \\ 0 \end{bmatrix} w^1 \\ \begin{pmatrix} y^1 \\ z^1 \end{pmatrix} = \begin{bmatrix} C^1 & 0 \\ M^1 & 0 \end{bmatrix} \begin{pmatrix} x^1 \\ \hat{z}^t \end{pmatrix} \end{cases} \qquad (4.113)$$

Now recall the definition in (4.29) to write

$$\begin{pmatrix} y^o_o \\ z \end{pmatrix} = \begin{pmatrix} y^1_1 \\ z \end{pmatrix} + \begin{pmatrix} \tilde{y}^1_1 \\ \tilde{z} \end{pmatrix} = \begin{bmatrix} C^1 \\ M^1 \end{bmatrix} x^1 + \begin{pmatrix} \tilde{y}^1 \\ \tilde{z}^1 \end{pmatrix} \qquad (4.114)$$

Substituting this in to (4.113) yields, using also (4.29),

$$\begin{pmatrix} \dot{x}^1 \\ \dot{\hat{z}}^t \end{pmatrix} = \begin{bmatrix} A^1 + B^1 GT_1 M^1 & B^1 GT_2 \\ \hline [d_{21} + B_2 GT_1] M^1 & d_{22} + B_2 GT_2 \end{bmatrix} \begin{pmatrix} x^1 \\ \hat{z}^t \end{pmatrix} + \begin{bmatrix} B^1 GT_1 \\ d_{21} + B_2 GT_1 \end{bmatrix} \tilde{z}^1 + \begin{bmatrix} I \\ 0 \end{bmatrix} w^1$$

$$\begin{pmatrix} y^o \\ z^o \end{pmatrix} = \begin{bmatrix} C^1 & 0 \\ M^1 & 0 \end{bmatrix} \begin{pmatrix} x^1 \\ \hat{z}^t \end{pmatrix} + \begin{pmatrix} \tilde{y}^1 \\ \tilde{z}^1 \end{pmatrix} \qquad (4.115)$$

It proves convenient to transform these coordinates to the coordinates $\begin{pmatrix} x^1 \\ \tilde{z}^t \end{pmatrix}$ where $\tilde{z}^t$ is defined by (4.82). Accomplish this by using (4.93),

(4.94) and (4.97) to get the simpler form for (4.115)

$$\begin{pmatrix} \dot{x}^1 \\ \dot{\tilde{z}}^t \end{pmatrix} = \left[\begin{array}{c|c} A^1 + B^1 GT_1 M^1 & -B^1 GT_z \\ \hline 0 & d_{22} \end{array} \right] \begin{pmatrix} x^1 \\ \tilde{z}^t \end{pmatrix}$$

$$\underbrace{\begin{bmatrix} I \\ 0 \end{bmatrix} w^1 + \begin{bmatrix} B^1 GT_1 \\ 0 \end{bmatrix} \tilde{z}^1}_{(a)} + \underbrace{\begin{bmatrix} B^1 GT_2 \Gamma \\ 0 \end{bmatrix} x^2 + \begin{bmatrix} 0 \\ -d_{21} \end{bmatrix} \tilde{z}^2}_{(b)} \qquad (4.116)$$

$$\begin{pmatrix} y^o \\ z^o \end{pmatrix} \quad \begin{bmatrix} C^1 & 0 \\ M^1 & 0 \end{bmatrix} \begin{pmatrix} x^1 \\ \tilde{z}^t \end{pmatrix} + \begin{pmatrix} \tilde{y}^1 \\ \tilde{z}^1 \end{pmatrix}$$

Notice that the system (4.116) is driven by the disturbance w^1 assumed for model $\mathscr{A}_1$ and the measurement residual $\tilde{z}^1 \triangleq z^o - z^1$, associated with model $\mathscr{A}_1$, and, in addition, the state $x^2(t)$ and measurement residual $\tilde{z}^2 \triangleq z^o - z^2$, associated with model $\mathscr{A}_2$. In general, both models are inexact and $\tilde{z}^1(t) \neq 0$ and $\tilde{z}^2(t) \neq 0$. The residuals $\tilde{z}^1$, $\tilde{z}^2$, may, in fact, be characterized as random processes which might have been ignored in the control problem but which now are to be considered in establishing estimator gains. Other important observations concerning (4.116) are to be made in the special cases which follow.

Suppose in (4.116) it happens that

$$T_1 M^1 x^1(t) + T_2 \Gamma x^2(t) \equiv x^1(t) \qquad (4.117)$$

then (4.116) reduces to

$$\begin{pmatrix} \dot{x}^1 \\ \dot{\tilde{z}}^t \end{pmatrix} = \begin{bmatrix} A^1 + B^1 G & -B^1 GT_2 \\ 0 & d_{22} \end{bmatrix} \begin{pmatrix} x^1 \\ \tilde{z}^t \end{pmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} w^1 + \begin{bmatrix} B^1 GT_1 \\ 0 \end{bmatrix} \tilde{z}^1 + \begin{bmatrix} 0 \\ -d_{21} \end{bmatrix} \tilde{z}^2$$

$$y^1 = [C^1 \quad 0] \begin{pmatrix} x^1 \\ \tilde{z}^t \end{pmatrix} \qquad\qquad (4.118)$$

$$z^1 = [M^1, \ 0] \begin{pmatrix} x^1 \\ \tilde{z}^t \end{pmatrix}$$

The nontrivial conditions under which (4.117) holds is that the ("evaluation") model $\mathscr{A}_1$ is identical to the ("design") model $\mathscr{A}_2$. That is $(A^1, B^1, C^1, M^1) = (A^2, B^2, C^2, M^2)$, since in that event (4.117) becomes

$$[T_1 M^2 + T_2 \Gamma - I] \, x^2(t) \equiv 0$$

which is guaranteed by the estimator design criteria (4.89b) repeated here, (which ensured nonsingularity of the transformation T in (4.79))

$$T_1 M^2 + T_2 \Gamma = 1$$

Now in this event where $\mathscr{A}_1$ and $\mathscr{A}_2$ are identical, (4.118) is rewritten

$$\begin{pmatrix} \dot{x}^2 \\ \dot{\tilde{z}}^t \end{pmatrix} = \begin{bmatrix} A^2 + B^2 G & -B^2 GT_2 \\ 0 & d_{22} \end{bmatrix} \begin{pmatrix} x^2 \\ \tilde{z}^t \end{pmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} \tilde{e}^{21} + \begin{bmatrix} B^2 GT_1 \\ -d_{21} \end{bmatrix} \tilde{z}^2$$

$$\begin{pmatrix} y^o \\ z^o \end{pmatrix} = \begin{bmatrix} C^2 & 0 \\ M^2 & 0 \end{bmatrix} \begin{pmatrix} x^2 \\ \tilde{z}^t \end{pmatrix} + \begin{pmatrix} \tilde{y}^2 \\ \tilde{z}^2 \end{pmatrix} \qquad\qquad (4.119)$$

where $\tilde{e}^{21}(t)$ represents those effects not included (modeled) in the model error vector $\hat{e}(t)$ which became an integral part of the model $\mathscr{A}_2$, see (4.46b). From the definition (4.26),

$$e^{21}(t) = \hat{e}(t) + \tilde{e}^{21}(t)$$

91

Note that (4.119) is stable if and only if the estimator is stable and the closed loop $\mathcal{A}_2$ system is stable.

It is instructive to write the solution for the linear system (4.119),

$$
y^o(t) = [C^2 \ 0] \left\{ \Phi(t,t_o) \begin{pmatrix} x_o^2 \\ \tilde{z}_o^t \end{pmatrix} + \int_0^t \Phi(t,\sigma) \left[\begin{bmatrix} I \\ 0 \end{bmatrix} \tilde{e}^{21} + \begin{bmatrix} B^2 G T_1 \\ -d_{21} \end{bmatrix} \tilde{z}^2 \right] d\sigma \right\}
$$
$$
+ \ \tilde{y}^2(t) \tag{4.120}
$$

where for the constant linear system $\Phi(t,\sigma)$ is defined by

$$
\Phi(t,\sigma) = \exp \begin{bmatrix} A^2 + B^2 G & -B^2 G T_2 \\ 0 & d_{22} \end{bmatrix} (t-\sigma) \tag{4.121}
$$

the partitioned parts of which are

$$
\Phi_{11}(t,0) = L^{-1} \left\{ sI - \left[A^2 + B^2 G \right] \right\}^{-1}
$$

$$
\Phi_{12}(t,0) = L^{-1} \left\{ \left[sI - [A^2 + B^2 G] \right]^{-1} [B^2 G T_2][sI - d_{22}]^{-1} \right\}
$$

$$
\Phi_{21}(t,0) = 0
$$

$$
\Phi_{22}(t,0) = L^{-1}[sI - d_{22}]^{-1}
$$

where L^{-1} denotes inverse Laplace transform and s denotes the complex variable. Using these partitioned forms rewrite (4.120) as

$$
y^o(t) = C^2 \left[\Phi_{11}(t,0) \ x_o^2 + \Phi_{12} \ \tilde{z}_o^t \right] + \int_0^t C^2 \Phi_{11}(t,\sigma) \tilde{e}^{21} \ d\sigma
$$

$$
+ \int_0^t C^2 \left[\Phi_{11}(t,\sigma) B^2 G T_1 - \Phi_{12}(t,\sigma) d_{21} \right] \tilde{z}^2 d\sigma + \tilde{y}^2(t) \tag{4.122}
$$

Now we note that the zero state response of the output with respect to the relative model errors in the plant, $\tilde{e}^{21}$, is independent of

92

estimator dynamics. Thus, these errors are transfered to the output via the transfer functions

$$\frac{y^o_{i}(s)}{\tilde{e}^{21}_{j}(s)} = \left\{ C^2 \left[sI - (A^2 + B^2 G) \right]^{-1} \right\}_{ij} \qquad (4.123)$$

Likewise it can be shown by taking the Laplace transform of (4.122) that the measurement prediction errors, $\tilde{z}^2$, propagate to the output, y^o, according to the transfer functions

$$\frac{y^o_{i}(s)}{\tilde{z}^2_{j}(s)} = \left\{ C^2 \left[sI - (A^2 + B^2 G) \right]^{-1} B^2 G \left[T_1 - T_2 (sI - d_{22})^{-1} a_{21} \right] \right\}_{ij} \qquad (4.124)$$

which transfer function possesses roots associated with the optimal control problem (due to the first term in (4.124)) and certain additional eigenvalues due to the estimator (but not equal to those of the estimator). Further insight is obtained by substituting the specific choices of the matrices A^2, B^2, C^2, and G defined by (4.46b) and (4.57).

The closed loop system matrix from the _control_ problem is, using (4.46b),

$$A^2 + B^2 G = \begin{bmatrix} A^3 & P \\ 0 & D \end{bmatrix} + \begin{bmatrix} B^3 \\ 0 \end{bmatrix} [G_1 \ G_2]$$

$$= \begin{bmatrix} A^3 + B^3 G_1 & P + B^3 G_2 \\ 0 & D \end{bmatrix} \qquad (4.125)$$

then since $C^2 = [C^3, 0]$, (4.123) becomes

$$\frac{y^o_{i}(s)}{\tilde{e}^{21}_{j}(s)} = \left\{ C^3 [sI - (A^3 + B^3 G_1)]^{-1} \left[I, + (P + B^3 G_2)(sI - D)^{-1} \right] \right\}_{ij} \qquad (4.126)$$

Now the relative error vector $\tilde{e}^{21}$ has the structure

$$\tilde{e}^{21} = \begin{pmatrix} \tilde{e}_1^{21} \\ \tilde{e}_2^{21} \end{pmatrix} \tag{4.127}$$

where $\tilde{e}_1^{21}$ and $\tilde{e}_2^{21}$ respectively force the plant $\mathscr{A}_3$ and the error system which is augmented to model $\mathscr{A}_3$. (The composite system is called model $\mathscr{A}_2$).

$$\dot{x}^3 = A^3 x^3 + B^3 u^o + e^{31} \tag{4.128}$$

But from (4.26)

$$e^{31} = \hat{e}^{31} + \tilde{e}^{31} \tag{4.129}$$

Now the particular choice for $\hat{e}^{31}$ in section 4.2 was from (4.43), (4.44)

$$\begin{aligned} \hat{e}^{31} &= P\gamma \\ \dot{\gamma} &= D\gamma \end{aligned} \tag{4.130}$$

Any errors associated with model $\mathscr{A}_2$ have the form

$$e^{21} \begin{pmatrix} \hat{e}^{31} \\ 0 \end{pmatrix} = \begin{pmatrix} \tilde{e}_1^{21} \\ \tilde{e}_2^{21} \end{pmatrix} = \begin{pmatrix} \hat{e}^{31} \\ 0 \end{pmatrix} + \begin{pmatrix} \tilde{e}^{31} \\ \tilde{e}_2^{21} \end{pmatrix} \tag{4.131}$$

since

$$\hat{e}^{21} = \begin{pmatrix} \hat{e}^{31} \\ 0 \end{pmatrix}$$

and (4.128) becomes

$$\dot{x}^2 = \begin{pmatrix} \dot{x}^3 \\ \dot{\gamma} \end{pmatrix} \begin{bmatrix} A^3 & P \\ 0 & D \end{bmatrix} \begin{pmatrix} x^3 \\ \gamma \end{pmatrix} + \begin{pmatrix} \tilde{e}_1^{21} \\ \tilde{e}_2^{21} \end{pmatrix} \tag{4.132}$$

where $\tilde{e}_1^{21} = \tilde{e}^{31}$. But from the definitions (4.27) then the error in the output (4.43) of the system (4.44) is

94

$$\tilde{e}_1^{21} \triangleq \int_0^t e^{D(t-\sigma)} \tilde{e}_2^{21} d\sigma = \tilde{e}^{31} \qquad (4.133)$$

Now from (4.126), (4.131) and (4.133) we have

$$y^o(s) = C^3 \left[sI - (A^3 + B^3 G_1) \right]^{-1} \left[I + (P + B^3 G_2) \right] [sI - D]^{-1} \tilde{e}_2^{21}(s) \qquad (4.134)$$

The characteristic equation for this transfer function will have the roots of the optimally controlled "error free" system $\mathcal{A}_3$, and in addition, the roots of the error system. The estimator dynamics do not enter into the transfer of relative errors, $\tilde{e}_2^{21}$, of the model error system, $\hat{e} = P\gamma$, $\dot{\gamma} = D\gamma$, to the output y^o, at least as far as the zero state response is concerned.

Now consider the substitution of the A^2, B^2, C^2 definitions into (4.124). The result is

$$y^o(s) = C^3 \left[sI - (A^3 + B^3 G_1) \right]^{-1} B^3 G \left[T_1 - T_2(sI - a_{22})^{-1} a_{21} \right] \tilde{z}^2(s) \qquad (4.135)$$

which relation involves, again, the closed loop dynamics of the optimally controlled "error free" system $\mathcal{A}_3$, and also the dynamics of the estimator (but not precisely the same eigenvalues).

Errors in the state equations of the error system, $\tilde{e}_2^{21}$, propagate to the output y^o in a manner governed by the eigenvalues of the optimal $\mathcal{A}_3$ "error free" system and the eigenvalues of the assumed modes of the error system, see (4.134). The propagation of errors in measurement prediction, $\tilde{z}^2$, to the output is governed by the eigenvalues of the optimal "error free" system $\mathcal{A}_3$ and the eigenvalues of the estimator, see (4.135).

Now the estimator gains generally should not be based exclusively upon considerations of either (4.134) or (4.135). The sum of the two, for given characterizations of $\tilde{z}^2(s)$ and $\tilde{e}_2^{21}(s)$, could be considered. However, difficulties arise when the error system model $\tilde{e}' = P\gamma$, $\gamma = D\gamma$ is unstable, a case that will prove desirable to consider in section 5.0. In such cases the transfer functions $y_i^o(s)/\tilde{e}_{2\,j}^{21}(s)$ are unstable. One possible remedy for this situation* is to choose the estimator gains in the best root mean square sense

$$\min_{T_{21}} \left[\lim_{T \to \infty} \frac{1}{T} \int_0^T y^o(t)^T Q \, y^o(t) dt \right]^{1/2}$$

where $y^o(t)$ is given by (4.120). Such a minimum exists if $y^o(t)$ is bounded. Another possibility is to select the estimator gains T_{21} to

$$\min_{T_{21}} \int_0^T y^o(t)^T Q \, y^o(t) dt$$

For any finite T, much larger than τ, the system characteristic time.

4.7 Selection of Estimator Gains On the Basis of Model $\mathcal{A}_1$.

A more appealing choice for the gains of the estimator than the model $\mathcal{A}_2$ referenced decisions of the previous section, is a choice

---

*After all, it is our own characterization of the error system that is unstable and not necessarily the actual "error system." Consider for example an actual error vector that is sinusoidal (resulting perhaps from a structural bending mode which was truncated from the design model). If the characteristic time of the controlled system is small compared to the period of this sinusoid, it might be quite reasonable to model such an error vector as a polynomial in time, since the state estimator continually updates the initial conditions associated with the error system rapidly compared to a period of the sine wave. In this way a bounded function is (piecewise) modeled as the output of an unstable system.

which involves an evaluation model ($\mathcal{A}_1$) which is better than the control design model ($\mathcal{A}_2$). To appreciate this point consider that it is usually the effects not included in the model upon which the controller design is based, which serve to limit the overall performance of the physical system, (which is under the influence of such a controller). The following procedure allows us to consider plant and measurement errors which were neglected in the previous modeling problems (and therefore in the estimator design). The flexibility offered by this "second chance" is as follows. To include all such (possibly correlated) model errors in the original statement of the problem causes the order of the estimator (and therefore of the controller) to be large. To neglect such disturbances all together limits the performance of the overall physical system plus controller, perhaps unacceptably.

To illustrate this point consider a completely deterministic design. That is, the model errors defined for the model are completely deterministic, as is suggested in the previous sections of this chapter and in the "external disturbance accommodation" work of Johnson.[3,4] But on the basis of the deterministic design there is no natural "bandwidth-limiting" mechanism. A requirement of asymptotic stability of the estimator design moves the eigenvalues "to the left of the imaginary axis" of the complex plane. The requirement that the estimator be faster than the system in the sense of placing eigenvalues "to the left of a vertical line $s = -1/\tau$," where $\tau > 0$ is a specified number (such as the characteristic time of the system, see definition (4.66)) are used by some (see section 4.0 of this paper

and Johnson,[3] page 20.) There are not, however, any known procedures

in the deterministic theory which naturally force the eigenvalues to

"lie to the right" of some line, thereby limiting the spectral radius

$(r[\bar{A}] \triangleq \lambda_{max} [\bar{A}]$ of the closed loop system matrix $\bar{A}$, and limiting the

amplification of system modeling errors.) In an engineering sense it

is precisely the feature of the stochastic optimization theory which

limits the spectral radius of the estimator (actually, by specifying

the poles uniquely, in the case of the Kalman filter), that is so

helpful. It is not the lack of uniqueness in the gains of the

Luengerger observer that invites poor pole placement, but the lack of

upper and lower bounds. The reason of course, that they are lacking

is that there is nothing in the model itself to warn of the limits of

its fidelity. There is associated with each model, $\mathcal{A}_j$, a finite

region in the complex plane, R_j, whose shape (and even connectedness)

is unknown, but which represents a region of "model credibility" (or

"spectrum of control authority") in the sense that closed loop system
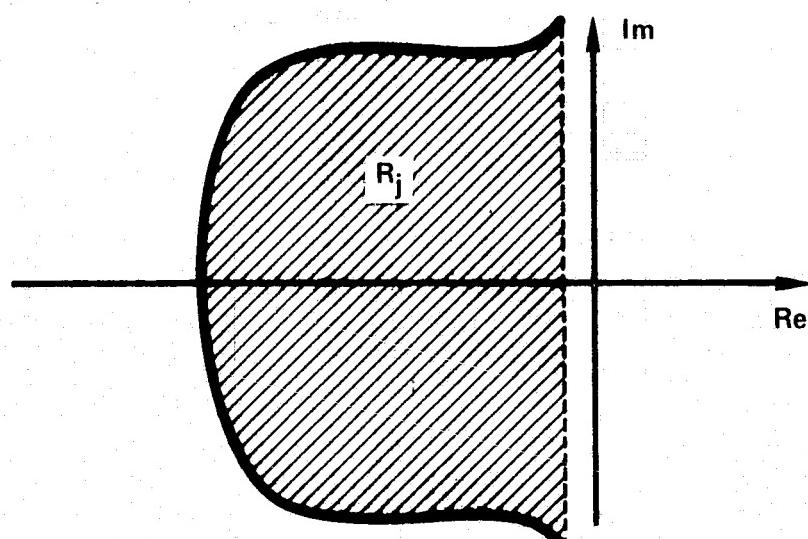
SPECTRUM OF CONTROL AUTHORITY

poles placed outside this region cause the closed loop system performance to "disagree" with that performance predicted on the basis of the model $\mathscr{A}_j$. Such a set, R_j is a "fuzzy" set in the sense that the boundaries cannot be precisely defined. Moreover the information to establish the boundaries of R_j is not continued within the model itself. These various reasons prompt us to investigate the possibility of designing controllers on the basis of more than one model. These "multi-model controller" designs have the advantage of the lower order that the minimal order Luenberger observes affords, together with the band limiting qualities afforded by the consideration of a second and better model.

Returning to the task of estimator gain evaluation on the basis of model $\mathscr{A}_1$, the equation of interest is (4.115), rewritten here in the compact form,

$$\dot{x} = Ax + G_z z^1 + G_w w^1 \qquad (4.136)$$
$$y^o = cx + y^1$$

where we may take X_o, $\tilde{z}^1$, w^1, $\tilde{y}^1$ to be either random or deterministic in character, and where

$$A \triangleq \left[\begin{array}{c|c} A^1 + B^1 GT_1 M^1 & B^1 GT_2 \\ \hline [d_{21} + B_2 GT_1] M^1 & d_{22} + B_2 GT_2 \end{array}\right]$$

$$G = \left[\begin{array}{c} B^1 GT_1 \\ d_{21} + B_2 GT_1 \end{array}\right] \quad , \quad G_w = \left[\begin{array}{c} I \\ 0 \end{array}\right]$$

$$C \triangleq [c^1, \; 0]$$

$$X \triangleq \begin{pmatrix} x^1 \\ \hat{z}^t \end{pmatrix}$$

99

We wish now to solve the problem

$$J^*_{(T_{21})} = \min_{T_{21}} E\left\{\frac{1}{T} \int_0^T y^{oT}(t) \, Q \, y^o(t) dt\right\} \tag{4.137}$$

for finite T (and also consider the limit as $T \to \infty$). Now if we

consider $G_d w_d$ to be the deterministic term forcing the linear plant

and $G_r w_r$ the zero mean white noise random process whose covariance is

known, and $\widetilde{y}^1$ is a zero mean white noise random process whose covar-

iance is known, then we can write (4.136) as

$$\dot{x} = Ax + G_d W_d + G_r W_r$$
$$y^o = Cx + \widetilde{y}^1 \tag{4.138}$$

where the specific definition of $G_d W_d$ and $G_r W_r$ will depend upon the

choice of $\widetilde{z}^1$ and w^1 characterizations in (4.136). From Appendix B,

Equation (B.10) we can immediately write (4.137) as

$$J^*(T_{21}) = \min_{T_{21}} \frac{1}{T} \, \text{tr} \, C^T Q C \int_0^T \left(\sum + \bar{x} \, \bar{x}^T\right) dt \tag{4.139}$$

where $\sum$ obeys

$$\dot{\sum} = A\sum + \sum A^T + G_r Q_r G_r^T \qquad \sum(0) = \sum_o = E[x_o x_o^T] \tag{4.140}$$

and $\bar{x}$ is defined as

$$\bar{x}(t) \triangleq \int_0^t \Phi(t,\sigma) G_d W_d d\sigma \tag{4.141}$$

where Φ is the state transition matrix for A, and satisfies

$$\dot{\Phi}(t,\sigma) = A \, \Phi(t,\sigma) \qquad \Phi(\sigma,\sigma) = I$$

Thus, it is not necessary to restrict the original model error vector

$\hat{e}^{21}(t)$ to a deterministic description. It is however, convenient to

do so, and no generality is lost* when one considers that the estimator

can be designed with full dimension (of order n_2) or with reduced

dimension (of order n_1, $n_2 - \ell \leq n \leq n_2$) and that in the full dimensional

case the gains may be selected for "optimality" in the stochastic

sense,** or for "stability" in the deterministic view. In the reduced

order estimator the gains may still be selected for stochastic opti-

mality in a limiting sense, corresponding to a Kalman filter with

vanishing measurement noise (see Kwatny[56]). Thus, it is convenient

and not restrictive (given any constraints on the order of the esti-

mator) to consider the deterministic design. In this view of the

model error description, the results of this research can be useful

to obtain a model, and a design, for systems subjected to correlated

disturbances. The modeling of correlated processes is an underrated

task.

It is well appreciated in stochastic problems that the condi-

tional expectation and more specifically the Kalman filter, is a

"biased" estimator if there are correlated disturbances (or "model

error vector") which have been modeled as white noise. It is common

to provide the simplest Gauss Markov model for such a disturbance, a

"bias" model (an unknown constant plus white noise, $\dot{\gamma} = 0 + w(t)$).

---

*Assuming Gauss Markov models of correlated disturbances are appro-
priately given by the deterministic structure given in (4.43),
(4.44) with white noise added to (4.44).

**If only the first and second moments (the means and the covariance
kernels) of the random processes, are known, then the Kalman filter
is the best linear estimate for a large class of optimization cri-
teria (including minimum variance), and this is true regardless of
the statistical properties of the random variables (see Theorem 5.3,
p. 166 of Reference 38).

Recall from section 4.2.2 that the modes present in the actual model error vector $e^{21}(t)$ which are important to model in the model error vector approximation, $\hat{e}(t)$, are those which "appear to be present in an interval τ" (τ is called in section 4 the characteristic time of the system). Therefore, the condition under which bias models are always sufficient as a characterization of a correlated disturbance is that the system characteristic time approach zero (corresponding to a controlled system with an infinite "bandwidth", to further abuse a common word in control engineering). It is perhaps unfortunate that in the vocabulary of the stochastic control literature the words "biased" and "unbiased" are used to describe estimators, when in fact the difference between the mean of the state and the mean of the estimate of the state is not at all constant. Nor should one be led to believe that the "biased" estimator can be made "unbiased" by the incorporation of "bias" disturbance models in the estimator (filter) design problem. The next simplest class of models for error vectors is the polynomial model (of which class the bias model is a member), and section 5.0 will treat this class of error vectors. The significance of that procedure in section 5.0 to stochastic problems is that the order of model for the correlated disturbance (error vector) is related to the characteristic time of the system (which time might have also been labeled the "correlation time"). This makes sense because white noise models are appropriate only for those processes whose correlation times are small compared to the system characteristic time. In this sense, section 5.0 can be viewed as a procedure to

determine that model of the system including disturbance and
measurement processes which can be "appropriately" driven by white
noise.

To summarize section 4.0 it is helpful to consider Figure 4-2
where the controller design events are summarized. The hierarchy of
models indicated in Figure 4-2 is utilized in the following way. The
physical system $\mathscr{A}_0$ is committed to a mathematical description, $\mathscr{A}_1'$,
which represents the physical system sufficiently closely to serve as
a final evaluation model (i.e., via computer simulation) to select
between candidate controller designs, (in the event the physical
system, $\mathscr{A}_0$ is not used for that purpose). The model $\mathscr{A}_1$ is a
linearized model whose order exceeds that acceptable for on-line state
estimation.

Model $\mathscr{A}_3$ is a truncation of $\mathscr{A}_1$ (whether actually constructed in
this way or not) and some errors between $\mathscr{A}_1$ and $\mathscr{A}_3$ are noted (i.e.
the specification of $\hat{e}^{31}$ by "assumed modes" (4.43), (4.44)). A con-
troller design on the basis of $\mathscr{A}_3$ is likely to be sensitive to the
modeling errors. Model $\mathscr{A}_2$ is therefore introduced, and is composed
of the model error system (4.43), (4.44) augmented to the truncated
model $\mathscr{A}_3$. The order of $\mathscr{A}_2$ is acceptable (actually by definition of
$\mathscr{A}_3$, $\mathscr{A}_2$) for controller design, and the structure of the controller
is established on the basis of the model $\mathscr{A}_2$. Now, the gains of the
estimator, T_{21}, may be selected on the basis of model $\mathscr{A}_2$, or, on the
basis of the better model $\mathscr{A}_1$. This latter "tri-model" design pro-
cedure (refer to the "design triangle" in Figure 4.2) is more difficult
computationally than conventional Kalman filter or Luenberger observer

103
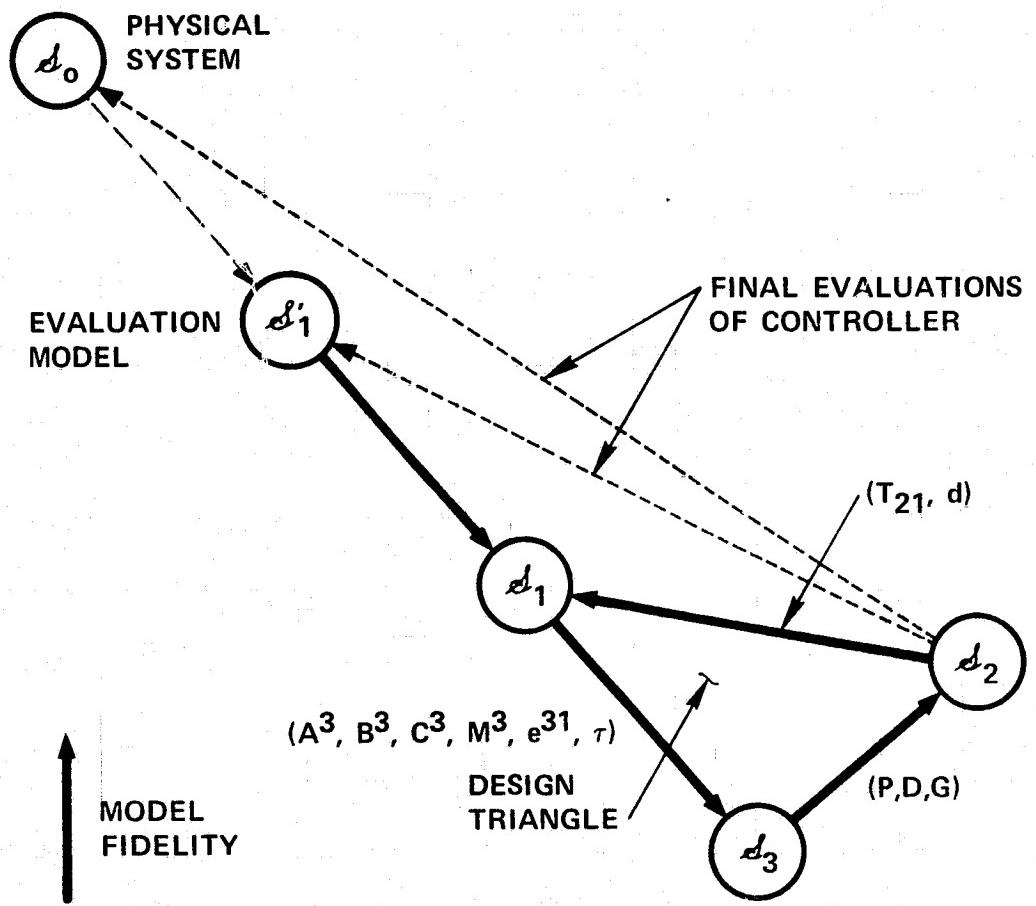
Figure 4-2. A Tri-Model Design Procedure.

designs (which are based upon a single system model which is assumed perfectly accurate) but has the following advantages. The reduced order Luenberger estimator has no natural "band limiting" features[*] unless white noise is considered present in the plant, in which case the estimator design becomes a Kalman filter in the limiting case of vanishing measurement noise. By considering the tri-model design procedure the estimator gains can be influenced by the presence of both (white) noisy characterizations of relative model error vectors $\hat{e}^{21}$ (which play the role of plant noise) and measurement residuals $\tilde{z}^1$ (measurement noise). The following section 5.0 helps to further systematize the choice of the model error system characterization (that is, the selection of $P(\tau,d)$ and $D(\tau,d)$).

---

[*] The eigenvalues can fall arbitrarily far to the left in the complex plane.

5.0 Orthogonal Filters for Model Error Estimation (OFFMEE)

In section 4.2.2 it has been suggested that the approximation of the model error vector $e(t)$ by $\hat{e}(t)$ be done with closed loop operation in mind so that state estimation can be applied to the differential equations the error is assumed to obey. Thus, $\hat{e}$ was chosen as a least squares approximation to $e(t)$ yielding the matrix of coefficients P given by (4.41) and the matrix D as determined to satisfy (4.17), where the $f_i(t)$ are a selected set of approximating functions. In this section the $f_i(t)$ are selected to be orthogonal functions with respect to the weight $g(t)$ on the interval $[0,\tau]$, see (4.41). Toward this end we introduce differential equations (for the error system) which orthogonal functions are known to obey. We will restrict our attention to the class of Jacobi polynomials which include via change of variable, Fourier, Legendre, Chebyshev and other polynomials. Before specializing to these cases, it is helpful to first develop a model error system of the form

$$\hat{e} = P\gamma$$
$$\dot{\gamma} = D\gamma$$

(5.1)

which is capable of generating the more general Jacobi polynomials. It will be convenient to normalize time to vary over the interval $[-1,1]$ instead of the interval $(0,\tau]$. Such a change of variable is accomplished by the definition,

$$\sigma \triangleq 2\frac{t}{\tau} - 1$$

(5.2)

Jacobi polynomials $J_i^{(\alpha,\beta)}(\sigma)$ are determined by the property of being orthogonal on the interval $-1 \leq \sigma \leq 1$, with respect to the weight

$g(\sigma) = (1-\sigma)^\alpha (1 + \sigma)^\beta$ and may be expressed in the form Reference 1, page 24,

$$J_i^{(\alpha,\beta)}(\sigma) = (1 - \sigma)^{-\alpha}(1+\sigma)^{-\beta}\frac{d^i}{dr^i}\left[(1-\sigma)^{\alpha+1}(1+\sigma)^{\beta+1}\right] \qquad (5.3)$$

Theorem 5.1:

The model error system which yields the least squares approximation of $\hat{e}(\sigma)$ to $e(\sigma)$ using Jacobi polynomials, $f_i(\sigma) = J_i^{(\alpha,\beta)}(\sigma)$, can be written, in terms of the normalized time, σ,

$$\hat{e}(\sigma) = P \, \gamma(\sigma)$$

$$\frac{d\gamma(\sigma)}{d\sigma} = D(\sigma) \, \gamma(\sigma) \qquad (5.4)$$

where

$$\gamma \in R^{zd}$$

$$\hat{e} \in R^{m}$$

$$P' = \int_{-1}^{1} g(\sigma)e(\sigma)f^T(\sigma)d\sigma \, \Lambda'$$

$$P = [P',0]$$

$$D = \begin{bmatrix} 0 & I_d \\ -\text{diag.}\begin{pmatrix} 0 \\ \dfrac{2+\alpha+\beta}{1-\sigma^2} \\ \dfrac{2(3+\alpha+\beta)}{1-\sigma^2} \\ \vdots \\ \dfrac{(d-1)(d+\alpha+\beta)}{1-\sigma^2} \end{pmatrix} & \left[\dfrac{-\beta+\alpha+(\alpha+\beta+2)\sigma}{1-\sigma^2}\right]I_d \end{bmatrix}_{2d \times 2d}$$

$$g(\sigma) = (1-\sigma)^{\alpha} (1+\sigma)^{\beta} , \qquad \alpha > -1 , \ \beta > -1$$

$$f^{T}(\sigma) = (J_{o}^{(\alpha,\beta)}(\sigma), \ J_{1}^{(\alpha,\beta)}(\sigma),\ldots,J_{d-1}^{(\alpha,\beta)}(\sigma))$$

where diag ($\cdot$) means a diagonal matrix with the ordered elements of ($\cdot$) on diagonal. This result, (5.4), is the first order form of the Sturm-Lionville, equations, Reference 62, page 60, and the eigenfunctions of (5.4) are Jacobi polynomials, $\gamma^{i}(\sigma) = f_{i}(\sigma) = J_{i}^{(\alpha,\beta)}(\sigma)$. Now, depending upon the choice for the parameters (α,β) the system (5.4) may generate Legendre polynomials ($\alpha = \beta = 0$), Chebyshev polynomials ($\alpha = \beta = -1/2$) and other polynomials of interest. It is also possible to select time invariant representations of (5.4) for the cases of interest which follow.

5.1 Chebyshev Error Systems

The special case of Theorem 5.1 when $\alpha = \beta$ yields ultraspherical polynomials of which class the Chebyshev polynomials are a member (specifically $\alpha = \beta = -1/2$). The Chebyshev polynomials are of primary interest in the remainder of our study because they include (by change of variable) the Fourier series. They also provide a simple and systematic extension to the "bias" models often used to model correlated disturbances in stochastic problems. Because of the good curve fitting properties of Chebyshev polynomials we introduce Theorem 5.2:

Suppose the model error system is modeled by

$$\hat{e}(\sigma) = P \ \gamma(\sigma)$$
$$\gamma'(\sigma) = \overline{D} \ \gamma(\sigma)$$

(5.5)

P is the set of Chebyshev-Fourier coefficients,

$$P = \int_{-1}^{1} g(\sigma)\, e(\sigma)\, f^{T}(\sigma)\, d\sigma\; \Lambda^{-1} \tag{5.6a}$$

$$\Lambda^{-1} = \frac{2}{\pi}\begin{bmatrix} 1/2 & 0 \\ 0 & I_{d-1} \end{bmatrix}$$

$$\overline{D} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 3 & 0 & 6 & 0 \\ & & & & \ddots \end{bmatrix}_{dxd} \tag{5.6b}$$

$$f^{T}(\sigma) = T^{T}(\sigma) = J_{o}^{(-1/2,-1/2)}(\sigma),\ J_{1}^{(-1/2-\,1/2)}(\sigma),\dots J_{d-1}^{(-1/2,-1/2)}(\sigma)$$

$$\tag{5.6c}$$

where $J_{1}^{(-1/2,-1/2)}(\sigma) \triangleq T_{i}(\sigma)$ are Chebyshev polynomials of the first

kind of degree i,

$$T_{1}(\sigma) \triangleq \cos(i\,\cos^{-1}\sigma)\quad,\qquad \sigma \triangleq \frac{2}{\tau}\,t - 1$$

Then, any solution of (5.5) satisfies

$$\hat{e}(\sigma) = P(\gamma_{o})\, T(\sigma) \tag{5.6d}$$

where

$$P(\gamma_{o}) \triangleq P\begin{bmatrix} \gamma_{0}^{T} & Q_{1} \\ \gamma_{0}^{T} & Q_{2} \\ & \vdots \\ \gamma_{0}^{T} & Q_{d} \end{bmatrix} \tag{5.6e}$$

(note: $\gamma(\sigma = -1) = \gamma(t = 0)$)

110

where

$$Q_1 = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & 0 \end{array}\right]_{dxd}$$

$$Q_2 = \left[\begin{array}{cc|c} 1 & 1 & \\ 1 & 0 & 0 \\ \hline & 0 & 0 \end{array}\right]_{dxd}$$

$$Q_3 = \left[\begin{array}{ccc|c} 3 & 4 & 1 & \\ 4 & 4 & 0 & 0 \\ 1 & 0 & 0 & \\ \hline & 0 & & 0 \end{array}\right]_{dxd}$$

$$\vdots$$

The proof follows from theorems A6 and A7 in Appendix A.

We wish to now interpret the performance of the state estimator for $\gamma(t)$, operating in the real time, $t = \frac{T}{2}(\sigma + 1)$.

Corollary:

Given the model $\mathscr{A}_3$

$$\mathscr{A}_3 \begin{cases} \dot{x}^3 = A^3 x^3 + B^3 u^o \\ z^3 = M^3 x^3 \end{cases}$$

(5.7a)

and the model $\mathscr{A}_2$

$$\mathscr{A}_2 \begin{cases} \dot{x}^3 = A^3 x^3 + B^3 u^o + \hat{e}(t) \\ z^2 = M^3 x^3 \\ \text{where} \\ \hat{e} = P \gamma \\ \dot{\gamma} = D \gamma \end{cases}$$

(5.7b)

111

$$D = \frac{2}{\tau} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 3 & 0 & 6 & 0 & \cdot \\ & & & & \ddots \end{bmatrix}_{dd} \tag{5.7c}$$

Then the difference between the physical measurements $z^0(t)$ and those predicted by model $\mathscr{A}_3$, $(z^3(t))$, is given by

$$z^0(t) - z^3(t) \triangleq z^3 = M^3 \, \Phi^3(t,t_0) \, P(\gamma_0) \, T(t) + \tilde{z}^2 \tag{5.7d}$$

where $\Phi^3(t,t_0)$ is defined by

$$\dot{\Phi}^3(t,t_0) = A^3 \, \Phi(t,t_0) \qquad\qquad \Phi(t_0, t_0) = I$$

and $\tilde{z}^2$ is defined by

$$\tilde{z}^2 \triangleq z^0 - z^2 \tag{5.7e}$$

and

$$T(t) \triangleq T(\sigma = 2/\tau \; t - 1) \tag{5.7f}$$

$\quad$ ($T(\sigma)$ defined by (5.6c))

and where $P(\gamma_0)$ is defined by (5.6e).

Proof: The proof follows immediately by construction of the solution of (5.7a), (5.7b), and the definitions in (4.29), and Theorem 5.2.

$\quad$ On the basis of Theorem 5.2 and its corollary, it is possible to interpret the performance of the state estimator (whose construction is based on model $\mathscr{A}_2$) in the following way. Assume that $\hat{e}^{31}(t)$ is a very good approximation to $e^{31}(t)$ so that, in terms of the definitions in (4.26), $\tilde{e}^{31} = 0$. Then $\tilde{z}^2 \equiv 0$ (by definition of e^{31} in section 4.1), and (5.7d) implies that the difference between the physical measurements $z^0(t)$ and those predicted by model $\mathscr{A}_3$ is equal to a linear

combination of Chebyshev polynomials, and furthermore the coefficients

of these polynomials are a function of the initial conditions, γ_0.

Recall from a discussion in section 4.5 that the presence of a nonzero

measurement residual, $z^0 - z^3$, causes the estimator to change its esti-

mate of the state vector from that predicted solely on the basis of

the model. Therefore observe from these comments and (5.7d) that if

the estimator is stable, then, in response to real measurements $z^0(t)$

the state estimator automatically selects those coefficients of

Chebyshev polynomials which cause improvement in the estimate of the

state vector $x^3(t)$ and the model error vector $e^{31}(t)$, such that those

measurements predicted by $z^3(t)$ are closer to those from the physical

system, $z^0(t)$. This is not "adaptive" curve fitting of the measure-

ment residual in the sense of changing parameters within a fixed model

structure, but an "adaptive" function has been served by increasing

the order of the model with the error system. The Chebyshev polynomials

used to approximate the model error, $e^{31}(t)$, which are automatically

selected by the state estimator for $\gamma(t)$, yield a least squares approx-

imation to the model error vector $e^{31}(t)$. Such an approximation of

e^{31} at any time, t, is based upon the present model error and the

prediction of model error τ units of time into the future. To see this

prediction feature, note the estimator design of section 4.5 and note

that the matrices P and D are functions of the prediction interval, τ.

This makes the inherent prediction feature of the estimator accurate

only over the selected interval τ. Selection of τ was related in

section 4.4 to a property of the closed loop system called the charac-

teristic time, see (4.66).

The estimator associated with the variables $\gamma(t)$ (noting that
$x^2 \triangleq \begin{pmatrix} x^3 \\ \gamma \end{pmatrix}$) can be thought of as an "orthogonal filter" to estimate

model error, since the model upon which that portion of the state

estimator associated with γ is based generates orthogonal functions.

The particular error system given by (5.5) is called a Chebyshev

error system. The design of "orthogonal filters" for model error esti-

mation (OFFMEE) is summarized in the specific steps of the next section.

Some further properties of the Chebyshev error system (5.7) need

to be mentioned. The eigenvalues of D are 0 repeated d times, and

this means that the error system model is unstable (see Appendix A).

The resulting closed loop system of a uni-model design is unstable.

That is, the analysis or "evaluation" of the controller with respect

to the same model, $\mathcal{d}_2$, upon which the controller design is also based,

indicates instability. To see this, refer to (4.11g) (further,

assume the model is perfect ($\tilde{z}^2$, $\tilde{e}^{21}$, $\tilde{y}^2$ all zero)). Then, using

(4.125), conclude that a subset of the eigenvalues of the closed loop

system (4.119) are those of D. Thus, on the basis of this "evaluation"

of the controller, some components of the state trajectory
$x^2(t) = \begin{pmatrix} x^3(t) \\ \gamma(t) \end{pmatrix}$ approach infinity as $t \to \infty$ from some nonzero initial

conditions. Two observations are pertinent to this situation. The

first has to do with the interpretations of this instability conclusion,

and the second has to do with the alternate suggestion for design of

the controller on the basis of more than one model of the system.

The concept of stability imposes limiting time, $t \to \infty$, arguments.

That is, from a given set of initial conditions, the initial conditions

must not change again $\forall$ t $\in$ [t$_o$, ∞). However, in the construction of
a model, we carefully noted that the model was capable of accurately
portraying the output of the physical systems only for a relatively
short time, and the particular time over which the model was argued to
be credible was the characteristic time, τ, defined in section 4.4.
It was specifically argued in the model error system construction of
section 4.0 that only modes of truncated states which appear to be
present over an interval τ need be modeled in the error system. For
example, suppose an uncontrollable mode of $\mathcal{A}_1$ were truncated from $\mathcal{A}_3$.
Let the mode have the effect of a damped sinusoidal forcing function
on the state equations. Let the effect of this truncated mode be
modeled (even exactly) over an interval τ as a second order polyno-
mial in time. Such modeling relies upon the state estimator to update
the initial conditions associated with the error system, (for accurate
modeling of the errors) but arguments of stability remove the (prac-
tical) circumstance of initial condition changes ($\forall$ t $\in$ [t$_o$,∞)) asso-
ciated with the assumed modes of the error system. Therefore, since
the model of the dynamical system is credible only over a relatively
small time, τ, (from a fixed initial condition) designing for stability
has less conceptual appeal than designing for performance of specific
variables as discussed next.

Suppose now, that all parameters are fixed by the controller
design except the estimator gains, T_{21} (see (4.110) and (4.105)).
Section 4.6 shows how the estimator may be stabilized. However, the
closed loop system, (4.115), (with respect to model $\mathcal{A}_2$) may not be
stable, from above arguments. If, instead of using $\mathcal{A}_2$ in (4.115),

we "evaluate" the controller with respect to a different model,

$\mathscr{A}_j \neq \mathscr{A}_\frown$, the system (4.115) might indicate stability. (i.e., any

model $\mathscr{A}_j$ for which $A^j + B^j \, GT_1 M^j$ is a stability matrix). Instead

of using stability arguments to fix the estimator gains, one might

choose to minimize the functional

$$J(T_{21}) = \frac{1}{T} \int_0^T y^T \, Qy \, dt \qquad (5.8)$$

(y is the output of (4.136) with $\mathscr{A}_1 = \mathscr{A}_2$) for finite T. Now again,

this functional is considered for a set of fixed initial conditions,

and in the limit as $T \to \infty$ we are faced with the same conclusions that

troubled us with the stability arguments: For fixed initial conditions,

$J(T_{21})$ might not provide a faithful prediction of the physical system

performance for large T. Making T equal to a fairly small multiple

of the control system's characteristic time, τ, might provide an answer

in closer agreement with the physical system. (i.e., $T = 20\tau$ or

$T = 10\tau$). It was mentioned early in this report that the performance

of controlled physical systems is usually limited by those effects not

modeled in the design of the control law. We may wish, therefore, to

choose T_{21} on the basis of (5.8) where y is the output of (4.136) (as

written, in terms of $\mathscr{A}_1$ model parameters).

The Case of Infinite Terminal Time and Chebyshev Error Systems.

The defining equation for L, Equation (4.53), can be used (in

the case of $T = \infty$) to produce a recursive solution for the elements

of L as follows. Since for the Chebyshev error system the matrix D

has a lower triangular form, the expression (from (4.53) with $\dot{L} \equiv 0$

corresponding to $T = \infty$ in (4.46a)),

116

$$L = \left[K^3 B^3 R^{-1} B^{3^T} - A^{3^T}\right]^{-1} \left[L\ D + K^3 P\right]$$

can be written

$$L_{ij} = \left\{\left[K^3 B^3 Z^{-1} B^{3^T} - A^{3^T}\right]^{-1}\right\}_{ik} \left[L_{k\ell}\ D_{\ell j} + [K^3 P]_{kj}\right]$$

where $\ell > j$ and the range of j is $j=1,2,\ldots d$. Thus, beginning with the dth column of the matrix L, denoted by L^d, where

$$L = [L^1, L^2, \ldots L^d] \quad , \quad P = [P^1, P^2, \ldots, P^d]$$

we have

$$L^d = \left[K^3 S^3 - A^{3^T}\right]^{-1} K^3\ P^d$$

due to the fact that the last column of D is zero, see (5.7c). Proceeding to write the columns of L proceeding right to left among the columns of L,

$$L^{d-1} = \left[K^3 S^3 - A^{3^T}\right]^{-1} \left(K^3 P^{d-2} + L^d\ D_{d,d-1}\right)$$

$$L^{d-2} = \left[K^3 S^3 - A^{3^T}\right]^{-1} \left(K^3\ P^{d-2} + L^{d-1}\ D_{d-1,d-2} + L^d D_{d,d-2}\right)$$

$$L^{d-3} = \left[K^3 S^3 - A^{3^T}\right]^{-1} \left(K^3 P^{d-3} + L^{d-2}\ D_{d-2,d-3} + L^{d-1} D_{d-1,d-3} + L^d D_{d,d-3}\right)$$

$$L^{d-4} = \left[K^3 S^3 - A^{3^T}\right]^{-1} \left(K^3 P^{d-4} + L^{d-3} D_{d-3,d-4} + L^{d-2} D_{d-2,d-4} + L^{d-1} D_{d-1,d-4} + L^d D_{d,d-4}\right)$$

$$\vdots$$

$$L^{d-j} = \left[K^3 S^3 - A^{3^T}\right]^{-1} \left(K^3 P^{d-j} + \sum_{k=0}^{j-1} L^{d-k} D_{d-k,d-j}\right) \quad , \quad j=0,1,\ldots,d-1$$

Thus, the above equations represent a solution for the columns of L

beginning with L^d as given, requiring only that the matrix D have a

lower triangular form (with zero diagonal elements). In the case of

the Chebyshev error system the matrix D has certain other zero entries

which may be omitted.

$$L^{d-j} = \left[K^3 S^3 - A^{3^T}\right]^{-1} \left(K^3 P^{d-j} + \sum_{k=1,3,5}^{j} D_{d-j+k,d-j} L^{d-j+k}\right) \quad j=0,1,2,\ldots,d-1$$

Lemma:

   If the square dxd matrix D has a lower triangular form, and the

nxn matrix S is nonsingular then the equation

$$SL + LD = F \qquad\qquad , D = \begin{bmatrix} 0 & & & & 0 \\ & 0 & 0 & & \\ & & 0 & 0 & \\ & & & \ddots & \ddots & \\ & & & & \ddots & 0 \end{bmatrix}$$

has the recursive solution which generates the d columns of L,

$$L^{d-j} = S^{-1}\left(F^{d-j} - \sum_{k=1,3}^{j}\left[D_{d-j+k,d-j} L^{d-j+k}\right]\right) \qquad j=0,1,\ldots d-1$$

beginning with the initial condition

$$L^d = S^{-1} F^d$$

5.2   Summary of a Tri-Model Design

      Procedure using OFFMEE

      In section 4.0 it was suggested that the controller design

involve more than one model. In section 5.0 it was suggested that the

model upon which the controller structure is based include an error

system such that the model errors are approximated with orthogonal

functions. The design of the controller based upon this augmented

system involved the "orthogonal filter" which estimates (and therefore reduces the impact of) the effects of the model errors which include external disturbances, truncated modes, and parameter errors. In this section we outline the step by step procedure to accomplish these design tasks.

Before outlining the steps it is helpful to review the total list of parameters which are to be determined, in order of their determination.

$\left.\begin{array}{c} A^3 \\ B^3 \\ C^3 \\ M^3 \end{array}\right\rbrace$ Parameters of the model $\mathscr{A}_3$

$$\left.\left\{\begin{array}{l} \dot{x}^3 = A^3 x^3 + B^3 u^o \\ \begin{pmatrix} y^3 \\ z^3 \end{pmatrix} = \begin{bmatrix} C^3 \\ M^3 \end{bmatrix} x^3 \end{array}\right\}\right\} \mathscr{A}_3$$

$\left.\begin{array}{c} Q \\ \\ \\ R \end{array}\right\rbrace$ weighting matrices for the optimal control problem

$$\min_{u^o} \int_0^T \left(y^{2^T} Q\, y^2 + u^{o^T} R\, u^o \right) dt$$

$K(A^3, B^3, C^3, Q, R)$ Riccati matrix for $\mathscr{A}_3$

$\tau(A^3, B^3, Q, R, K)$ Characteristic time of $\mathscr{A}_3$

$e^{31}(t)$ assumed "mode shape" of model error vector

$D(\tau, d)$ Matrix for synthetic modes of error system

$P(\tau, d)$ Chebyshev-Fourier coefficients of error system

$L(P, D, K, A^3, B^3, R) \Longrightarrow G(R, B^3, K, L).$ Feedback gains for model error

estimates

119

$\tilde{z}^1$ measurement errors in model $\mathcal{A}_1$ which are to be considered in estimator gain determination

$\tilde{e}^{21}$ model errors in $\mathcal{A}_2$ which are to be considered in estimator gain determination

$T_{21}(\tilde{z}^1, \tilde{e}^{21}, M^3$ all of above) Estimator gains, selected on the basis

of either 1) model $\mathcal{A}_2$

2) model $\mathcal{A}_1$

$d(J(\mathcal{A}_1))$ Order of the error system

The procedure is described for the case of infinite time solutions of the optimal control problem. There are special considerations to be made, such as imposing an artificial (large) finite value for terminal time, when existence of an infinite integral is questioned, and these are cases appropriately noted.

ASSUMPTIONS:

    1. controllability of (A^3, B^3)

    2. observability of (D,P), (A^3, C^3), (A^3, M^3)

    3. Q, R both positive definite

STEP I: Determine $\mathcal{A}_3$ (Reduce from $\mathcal{A}_1$ or write directly)

$$\boxed{\text{specify } (A^3, B^3, C^3, M^3)}$$

STEP II: Solve the error free problem.

$$\boxed{\text{specify } (Q,R) \text{ and determine } (K, G_1) \text{ from}}$$

$$0 = -KA^3 - A^{3^T}K + KB^3R^{-1}B^{3^T}K - C^{3^T}QC^3$$

$$G_1 = -R^{-1}B^{3^T}K \qquad (5.9)$$

STEP III: Solve for τ, the characteristic time of $\mathcal{d}_3$

<div style="text-align:center">| determine τ |</div>   from

$$\tau = \lambda_{max} \left\{ \left[KB^3 R^{-1} B^{3^T} K + C^{3^T} QC^3 \right]^{-1} K \right\}$$ (5.10)

STEP IV: Assume a characteristic shape for $e^{31}(t)$

<div style="text-align:center">| determine $e^{31}(t)$ |</div>   from

a) compute precisely that $e^{31}(t, x_o^3, x_o^1)$ corresponding

to $\mathcal{d}_1 \longrightarrow \mathcal{d}_3$, or

b) assume some shapes expected in $e^{31}(t)$ (damped sine-

waves, polynomials, etc.)

STEP V: Select orthogonal approximating functions $f_i(t)$ (i.e.,

Chebyshev, Legendre, Fourier, etc.)

<div style="text-align:center">| Specify D |</div>   from

$\dot{f} = D \ f$

For Chebyshev error systems

$$D(\tau,d) \triangleq \frac{2}{\tau} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 3 & 0 & 6 & 0 \\ & & & & \ddots \end{bmatrix}_{dxd}$$ (5.11)

STEP VI: Solve for the Chebyshev-Fourier coefficients, P.

<div style="text-align:center">| determine $P(\tau,d)$ |</div>   from

$$P = \int_{-1}^{1} g(\sigma) \, e^{31}(\sigma) \, f^T(\sigma) d\sigma \ \Lambda^{-1} \qquad\qquad (5.12)$$

where

$$\Lambda^{-1} = \left[\int_{-1}^{1} g(\sigma) f(\sigma) f^T(\sigma) d\sigma\right]^{-1} = \frac{2}{\pi} \begin{bmatrix} 1/2 & 0 \\ 0 & I_{d-1} \end{bmatrix}$$

$$\sigma \triangleq 2\frac{t}{\tau} - 1 \qquad f_i(\sigma) = T_i(\sigma) = J_i^{(-1/2,\,-1/2)}(\sigma) = \cos(i \cos^{-1}\sigma)$$

STEP VII: Solve for error system feedback gains

$$\boxed{\text{Determine} \quad G_2 = - R^{-1}B^{3^T}L} \qquad \text{from}^*$$

$$\left[KB^3R^{-1}B^{3^T} - A^{3^T}\right] L - LD = KP \qquad\qquad (5.13)$$

Selection of estimator gains may be based upon either of two options
(VIIIa, IXa) or (VIIIb)

(STEP VIIIa): Specify measurement errors, $\tilde{z}^1$, and relative model

errors, $\tilde{e}^{21}$, (which <u>are to be neglected</u> in designing

the <u>structure</u> of the estimator but which are <u>to be</u>

<u>considered</u> in the selection of estimator gains.)

$$\boxed{\text{Specify} \quad \tilde{z}^1(t) \,,\, \tilde{e}^{21}(t)} \qquad \text{from either}$$

a) deterministic: $\tilde{z}^1(t)$, $\tilde{e}^{21}(t)$

b) random: $E[\tilde{z}^1]$, $E(\tilde{z}^1 - \bar{\tilde{z}}^1)(\tilde{z}^1 - \bar{\tilde{z}}^1)^T$

$\qquad\qquad E[\tilde{e}^{21}]$, $E(\tilde{e}^{21} - \bar{\tilde{e}}^{21})(\tilde{e}^{21} - \bar{\tilde{e}}^{21})^T$

---

*In the infinite time problem the optimality of G_2 has not been
guaranteed but the existence of L has, provided (A^3, B^3, C^3) is a
controllable, observable model.

<u>(STEP IXa):</u>   Solve for estimator gains

$$\boxed{\text{determine } T_{21}}$$ from

$$\min_{T_{21}} \left\{ \frac{1}{T} \ \text{tr } C^T Q C \int_0^T \left(\Sigma + \bar{x} \, \bar{x}^T \right) dt \right\}$$

where Σ obeys

$$\dot{\Sigma} = A\Sigma + \Sigma A^T + G_r Q_r G_r^T \quad , \quad \Sigma(0) = E\left[x_o x_o^T \right] = \Sigma_1 \qquad (5.14)$$

and $\bar{x}$ is defined by

$$\bar{x}(t) = \int_0^t \Phi(t,\sigma) G_d W_d d\sigma \qquad (5.15)$$

where $\Phi(t,\sigma)$ satisfies

$$\dot{\Phi}(t,\sigma) = A \, \Phi(t,\sigma) \qquad\qquad \Phi(\sigma,\sigma) = I$$

and

$$A = \left[\begin{array}{c|c} A^j + B^j G T_1 \ M^j & B^j G T_2 \\ \hline [a_{21} + B_2 G T_1] M^j & a_{22} + B_2 G T_2 \end{array} \right] \qquad (5.16)$$

where T_1 and T_2 are defined by (4.102) and (4.103).

$$T_1 = \left[\begin{array}{c} \Gamma_{11}^{-1} \left[I - \Gamma_{12} \, T_{21} \right] \\ T_{21} \end{array} \right]$$

$$T_2 = \left[\begin{array}{c} - \Gamma_{11}^{-1} \ \Gamma_{12} \ T_{22} \\ T_{22} \end{array} \right] \qquad \text{(TAKE } T_{22} = I \text{ without loss of} \\ \text{generality)}$$

$$[\Gamma_{11}, \ \Gamma_{12}] \triangleq M^j$$

STEP VIIIb: Solve for estimator gains

$$\boxed{\text{determine } T_{21}}$$ from

$$T_{21} = - K_T \, G_2^T \, R_\xi^{-1} \qquad\qquad (5.17)$$

where

$$0 = -K_T F_2^T - F_2 K_T + K_T G_2^T R_\xi^{-1} G_2 K_T - Q_\xi^T Q_\xi \qquad (5.18)$$

$$F_2 \triangleq \begin{bmatrix} 0_{q\ell} & , & I_q \end{bmatrix} \begin{bmatrix} A_{n_3}^3 & P_{n_3 d} \\ 0_{d n_3} & D_d \end{bmatrix} \begin{bmatrix} -\Gamma_{11_\ell}^{-1} & \Gamma_{12_{\ell q}} \\ & I_q \end{bmatrix}^* \qquad (5.19)$$

$$G_2 \triangleq - M^2 \begin{bmatrix} A_{n_3}^3 & P_{n_3 d} \\ 0_{d n_3} & D_d \end{bmatrix} \begin{bmatrix} -\Gamma_{11_\ell}^{-1} & \Gamma_{12_{\ell q}} \\ & I_q \end{bmatrix}^* \qquad (5.20)$$

$$\begin{bmatrix} \Gamma_{11_\ell} & , & \Gamma_{12_{\ell q}} \end{bmatrix} \triangleq M^2_{\ell n_2} \qquad , \qquad \begin{pmatrix} n_2 \triangleq n_3 + d \\ q \triangleq n_2 - \ell \end{pmatrix}$$

R_ξ , Q_ξ any positive definite matrices.

STEP IX: $\boxed{\text{Specify } d}$ from

$$V^2 = \frac{1}{T} \int_0^T y^{2^T} Q \, y^2 \, dt \le V^0 \qquad (\text{iterate for } d = 0, 1, 2, \dots)$$

Several facts are summarized here concerning the options for estimator gain selection, T_{21}. (See further discussion in sections 4.5, 4.6).

---

*The subscripts on the matrices $[\cdot]_{ab}$ denote dimensions a x b, unless a = b, in which case only one subscript is used.

124

Consider the design of the estimator for model $\mathscr{A}_2$ (the order, n, of the estimator is then $n_2 - \ell \leq n \leq n_2$). It is helpful to keep in mind that the Luenberger observer has no natural "band limiting" feature which is imposed upon gain selection, while the Kalman filter has no natural lower limit on the gains.

1. If the order of the estimator is n_2, then it is possible to select the gains so that the estimator becomes the steady state Kalman filter. (Implying that the model error $\tilde{e}^{21}$ in (4.132) and the measurement residual error $\tilde{z}^2$ (see (4.29)) may both be considered to be white noise even though they are neglected altogether in the deterministic point of view which leads to the same estimator).

2. If the order of the estimator is $n_2 - \ell$, as outlined in section 4.5, then the gains may be selected in STEP VIIIb so that the estimator still becomes a limiting form of the Kalman filter corresponding to the case of vanishing measurement noise. (In this way the selection of T_{21} is influenced by white noise errors $\tilde{e}^{21}$ but cannot be influenced by the measurement residual errors, $\tilde{z}^2$).

3. If the order of the estimator is $n_2 - \ell$, as in section 4.5, and the estimator gains are selected as in STEPS VIIIa, IXa, then the presence of random errors in $\tilde{e}^{21}$ and the presence of random errors in $\tilde{z}^2$ both influence the selection of gains, in addition to any deterministic effects presumed present.

The following section illustrates these steps of the design procedure for a practical problem.

6.0 An OFFMEE Control Design for the Large Space Telescope (LST).

The control problem to be considered is that of the Large Space Telescope (LST) sponsored by the National Aeronautics and Space Administration (NASA). The LST is an unmanned 3 meter optical telescope designed for extra-galactic research, and diffraction-limited performance of the optical system must be guaranteed to assure feasibility of the mission. From the Rayleigh criteria two stars can be discerned if they are $\theta = 1.22 \frac{\lambda}{D}$ radians apart, where λ is the average wavelength of the visible spectrum of the stars and D is the diameter of the telescope. The average wavelength of the visible region to be used for LST design purposes is λ = 628.3 nanometers and the diffraction limit of a 3 meter telescope is therefore

$$\theta = 1.22 \frac{628.3 \times 10^{-4}}{3} \left(2.06 \times 10^5 \frac{\widehat{sec}}{rad}\right) \approx .05 \ \widehat{sec}$$

where $\widehat{sec}$ denotes seconds of arc. The control system for pointing the telescope must allow dynamic errors much smaller than this diffraction limit of .05 sec. Considering the body-pointing control system requirement to be .005 $\widehat{sec}$, the problem then is to control this rather flexible cylinder which measures 4m × 20m in such a way that the optical axis is not perturbed more than .005 $\widehat{sec}$, rms during an exposure interval of up to six hours.

The disturbance environment depicted in Figure 6-1 includes aerodynamic, magnetic, and gravity gradient torques, spurious torques from sensor noise, and vibrational "disturbances" from the flexible structure and from imperfect rotors in the particular choice of momentum
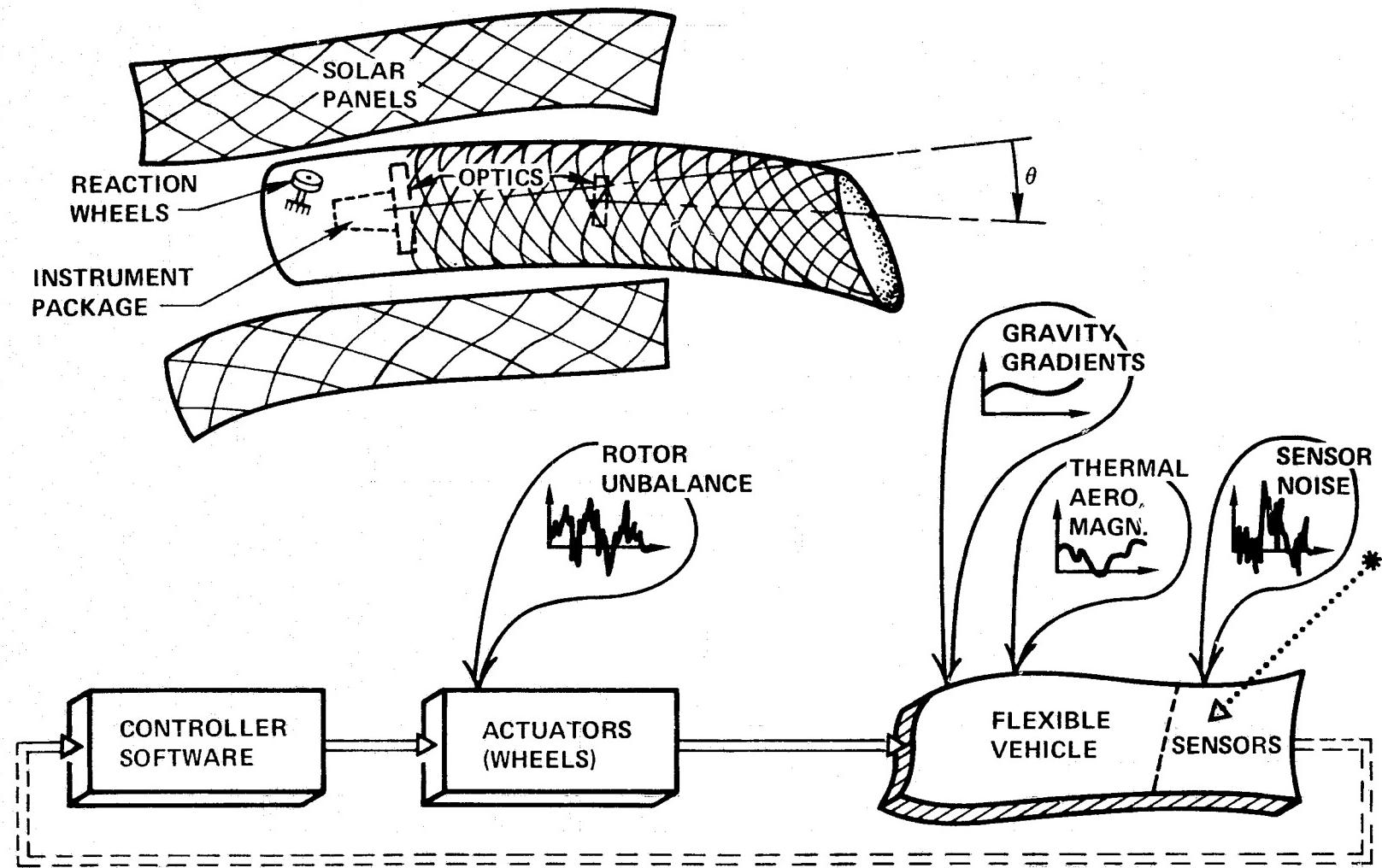
Figure 6-1. Large Space Telescope.

exchange controllers called reaction wheels. Modal analyses of the
flexible LST structure have identified structural modes from .1 hz
upward. There are a number of problems in strategy that immediately
beset the control designer: He must <u>develop a model</u> of the physical
system including "significant" dynamical effects. He must <u>design the</u>
<u>controller</u> to actively control "significant" modes of the total
dynamical system (including internal and external disturbing
effects). The interdependence of the modeling problem and control
problem is evident since one cannot usually know what dynamical modes
of the system must be actively controlled (are "significant") prior to
establishing a model. The model which is appropriate depends, in
turn, on what disturbing effects must be controlled. In practice
several iterations of models and control laws are usually tried. One
may avoid such iteration by working with the most general, or "con-
servative," model first (which is credible at "all" frequencies). In
practice, however, such a model might be extremely costly (in time and
resources), if not impossible, to construct. Moreover, if it were
used for control design, a complex controller may result which also
feeds back insignificant variables. Thus, the economic motivation of
the Minimal Controller Problem of section 2.0 comes to the fore:
Design the least complex controller that will satisfy the performance
requirements. From this motivation we reduce our large model $\mathcal{A}_1$ of
LST, derived in Appendix C, to a very simple model $\mathcal{A}_3$. We augment to
$\mathcal{A}_3$ an "error system" to partially compensate for the modeling errors
associated with this drastically reduced model.

The "line of sight," or the "optical axis" of the telescope can be displaced due to the effects of thermally induced displacements of the solar panels. This "disturbance" turns out to be significant and the effects are approximated in the derivation of $\mathcal{A}_1$ in Appendix C. The approximations of gravity gradient, magnetic, and aerodynamic disturbances are based upon a rigid spacecraft and are also summarized in Appendix C.

6.1 Discussion of The Model $\mathcal{A}_1$

In Appendix C a linearized model for LST is developed considering the effects of spacecraft flexibility, thermally induced appendage motions, gravity gradient, and aerodynamics. The model given by (C-77) is

$$\dot{x}^1 = A^1 x^1 + B^1 u^o + E^1 w^1$$

$$y^1 = C^1 x^1 \tag{6.1}$$

$$z^1 = M^1 x^1$$

where A^1, B^1, E^1, C^1, M^1, x^1, y^1, z^1, u^o, w^1 are specified by

$$x^1 \triangleq \begin{pmatrix} ^{23}\theta^{11} \\ ^{23}\theta^{1} \\ ^{23}\omega^{11} \\ ^{23}\omega' \\ ^{23}\dot{\omega}' \\ \chi \\ \dot{\chi} \\ h_s \\ i \end{pmatrix} , \quad y^1 \triangleq \begin{pmatrix} \theta_2 \\ \theta_3 \\ \dot{\theta}_3 \\ \ddot{\theta}_3 \end{pmatrix} , \quad z^1 \triangleq \begin{pmatrix} ^{23}\theta^{11} \\ ^{23}\omega^{11} \end{pmatrix}$$

$$u^o \triangleq \dot{h}' , \quad w^1 = \begin{pmatrix} \hat{T} \\ \bar{k} \end{pmatrix}$$

where

$$\hat{T} = \begin{pmatrix} T_o \\ \ddot{\Delta}(t) \end{pmatrix} \qquad \maltese \triangleq \begin{pmatrix} \theta \\ \eta \end{pmatrix}$$

and
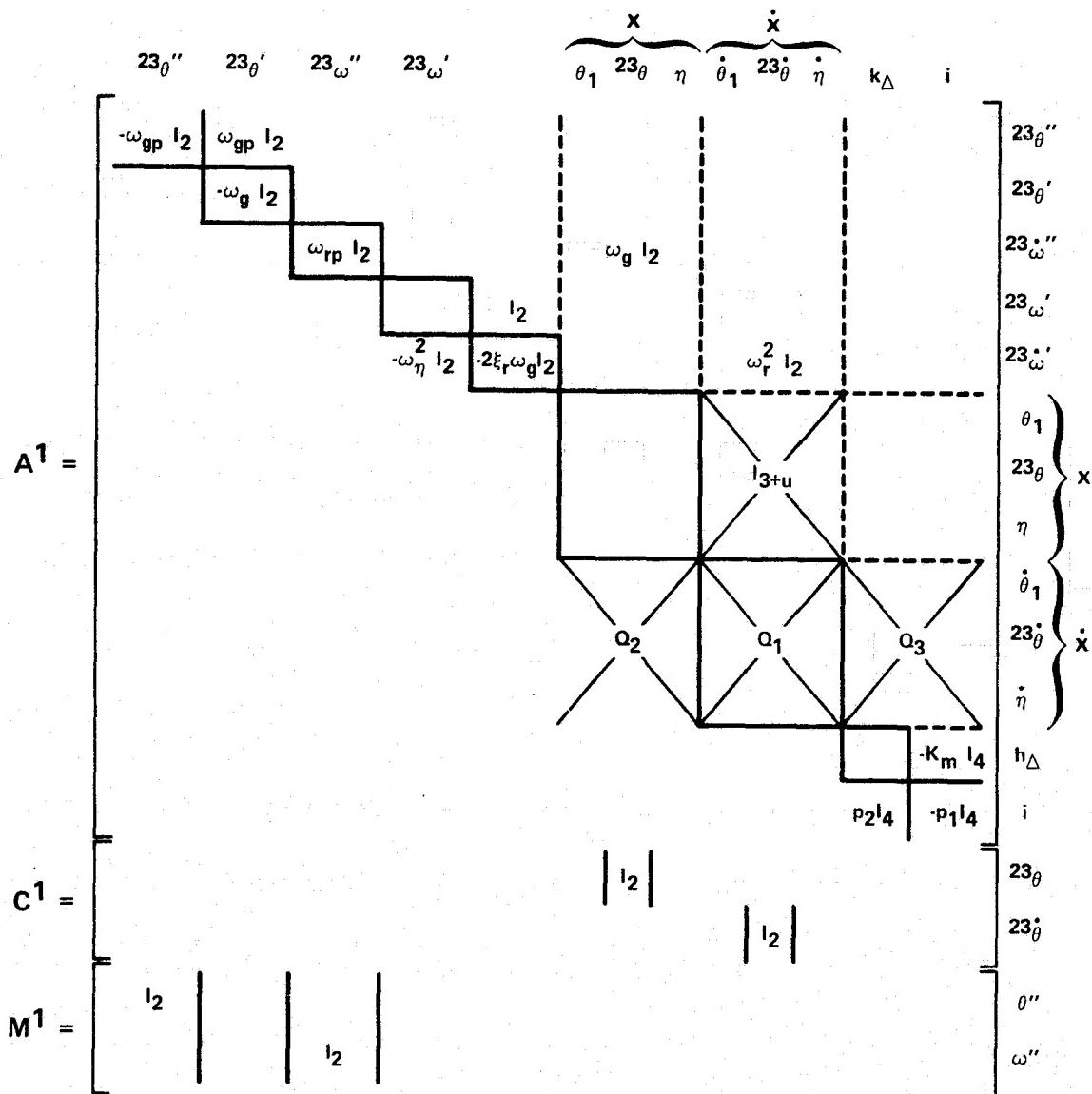
$$
B^1 = \begin{bmatrix} & u^o & \\ & 0 & \\ & & \\ & & \\ -P_3 & I_4 \end{bmatrix}
\qquad\qquad
E^1 = \begin{bmatrix} T & \bar{h} \\ & & \\ 0 & 0 \\ & & \\ \mathcal{D}_4 & 0 \\ & & \\ 0 & P_2 \end{bmatrix}
\begin{matrix}
\theta^{11} \\ \theta^1 \\ \omega'' \\ \omega' \\ \dot{\omega}' \\ \chi \\ \dot{\chi} \\ h_\Delta \\ i
\end{matrix}
$$

where empty partitions in matrices imply zero entries. See Appendix C for parameter definitions. Here, x^1 is a $(24 + 2\nu)$-vector where ν is the number of flexible appendage modes retained in the model $\mathcal{A}_1$ ($\nu = \dim n$). Of the 24 other variables, 4 are associated with rate gyros, 2 with position sensors, 2 with rate signal filtering, 2 with position signal filtering, 8 with reaction wheel control and 6 with the rigid body dynamics. Figure C-3 illustrates in block diagram form the general model. Now according to the STEPS outlined in section 5.2 we begin our OFFMEE design tasks.

6.2 Development of the model $\mathcal{A}_3$ (STEP I)

The pursuit of formal procedures for reducing the model from $\mathcal{A}_1$ to $\mathcal{A}_3$, shown as Step I in Figure 4-2 is reserved for further research. One attractive method for accomplishing such a reduction, given the desired order of $\mathcal{A}_3$, is presented in Reference 11 but the method requires manipulations of $\mathcal{A}_1$ which are costly owing to its large dimension. We will proceed to select a simple model for $\mathcal{A}_3$ directly

which is based upon ad hoc procedures in order to illustrate the procedures of sections 4.0 and 5.0.

It was argued in Appendix C that the cross product terms, $(S \, \overline{h})^{\sim} \dot{\theta}$, between controller momentum and vehicle rate were small, as were the non-secular terms from external disturbances, $G(t)\theta$. Equation (C-74) then becomes

$$\begin{pmatrix} \ddot{\theta} \\ \ddot{\eta} \end{pmatrix} = \hat{J}^{-1} \begin{bmatrix} 0 & 0 \\ 0 & -2\zeta\sigma \end{bmatrix} \begin{pmatrix} \dot{\theta} \\ \dot{\eta} \end{pmatrix} + \hat{J}^{-1} \begin{bmatrix} 0 & 0 \\ 0 & -\sigma^2 \end{bmatrix} \begin{pmatrix} \theta \\ \eta \end{pmatrix} + \hat{J}^{-1} \begin{bmatrix} I_3 \\ 0 \end{bmatrix} T_c$$

$$+ \hat{J}^{-1} \begin{bmatrix} I_3 & 0 \\ 0 & \Phi^T f_{\mathcal{M}} \end{bmatrix} \begin{pmatrix} T_o \\ \ddot{\Delta} \end{pmatrix} \qquad (6.2)$$

where the additional definition has been substituted, $^r SK_m i \triangleq T_c$, where T_c is taken as an approximation of the control torque. Noting the definition of $\hat{J}^{-1}$ in (C-75) this model becomes

$$\begin{pmatrix} \ddot{\theta} \\ \ddot{\eta} \end{pmatrix} = \begin{bmatrix} 0 & F^{-1} \Gamma\Phi \, 2\zeta\sigma \\ 0 & -[I + \Phi^T\Gamma^T F^{-1}\Gamma\Phi]2\zeta\sigma \end{bmatrix} \begin{pmatrix} \dot{\theta} \\ \dot{\eta} \end{pmatrix} + \begin{bmatrix} 0 & F^{-1} \Gamma\Phi \, \sigma^2 \\ 0 & -[I + \Phi^T\Gamma^T F^{-1}\Gamma\Phi]\sigma^2 \end{bmatrix} \begin{pmatrix} \theta \\ \eta \end{pmatrix}$$

$$+ \begin{bmatrix} F^{-1} \\ -\Phi^T\Gamma^T F^{-1} \end{bmatrix} T_c + \begin{bmatrix} F^{-1} & -F^{-1}\Gamma\Phi\Phi^T f_{\mathcal{M}} \\ -\Phi^T\Gamma^T F^{-1} & [I+\Phi^T\Gamma^T F^{-1}\Gamma\Phi]\Phi^T f_{\mathcal{M}} \end{bmatrix} \begin{pmatrix} T_o \\ \ddot{\Delta} \end{pmatrix} \quad (6.3)$$

where $F^{-1} \triangleq [J^* - \Gamma\Phi\Phi^T\Gamma^T]^{-1}$. (The Γ of this section should not be confused with Γ of earlier sections. See Appendix C for definitions). Now if we choose[*] to truncate all appendage modes from the model, we must write, in the spirit of section 4.0,

---

[*]Generally we might choose to keep some appendage modes within the design model and truncate others, depending among other things upon the uncertainty in the modal data.

$$\ddot{\theta} = F^{-1} T_c + e(t) \qquad (6.4)$$

where the model error vector e^{31} evolves from the error system

(regarding (6.2) as model $\mathscr{A}_1$ for our present purposes),

$$e = F^{-1} \Gamma \Phi (25\sigma \dot{\eta} + \sigma^2 \eta) + F^{-1}(T_o - \Gamma \Phi \Phi^T \, {}^f\!\mathscr{M} \ddot{\Delta})$$

$$\ddot{\eta} = - [I + \Phi^T \Gamma^T F^{-1} \Gamma \Phi][2\zeta\sigma \dot{\eta} + \sigma^2 \eta] - \Phi^T \Gamma^T F^{-1} T_c$$

$$- \Phi^T \Gamma^T F^{-1} T_o + [I + \Phi^T \Gamma^T F^{-1} \Gamma \Phi]\Phi^T \, {}^f\!\mathscr{M} \ddot{\Delta} \qquad (6.5)$$

If we further simplify the model to a single axis problem (principal

coordinates assumed) about axis $\underline{b}_2$, then

$$\ddot{\theta}_2 = \{F^{-1} T_c\}_2 + \underset{\sim}{e}_2^* \qquad (6.6)$$

$$\text{error} \begin{cases} e_2 = \{F^{-1}\Gamma\Phi(2\zeta\sigma \dot{\eta} + \sigma^2 \eta)\}_2 + \{F^{-1}(T_o - \Gamma\Phi\Phi^T \, {}^f\!\mathscr{M} \ddot{\Delta})\}_2 \\[2mm] \text{system} \; \ddot{\eta} = - \Phi^T \Gamma^T F^{-1}(T_o + T_c) + [I + \Phi^T \Gamma^T F^{-1} \Gamma \Phi]\Phi^T \, {}^f\!\mathscr{M} \ddot{\Delta} \end{cases} \qquad (6.7)$$

An alternate approach of approximating the truncated effects is
to identify a small parameter which multiplies the coupling terms on
the left hand side of Equation (C-73) and apply singular perturbation
theory in the solution of the control problem. The methods of[8] may
be applied in this event as discussed in section 3.3. Such an approach,
however, seeks to implement a control law which feeds back all the
states of the higher dimension of (C-73) and is an approximate solution
to a higher order problem. We wish instead to formulate a version of
the problem that is less exact (than (C-73) represents) and then solve
this approximate problem exactly. Toward this end we identify some

---

*For convenience, we drop the 31 superscript on the error vector.

134

properties of truncated effects and implement an approximation which evolves from an "error system" which is of smaller order than the truncated modes, η in (C-73).

Returning to our immediate task of constructing a simplified model, $\mathscr{A}_3$, we noted in the discussion of section 4.0 that the decision of how many modes to assign to the physical variables in $\mathscr{A}_3$ and how many modes to assume in the error system was not a straight-forward decision to make even for fixed dimension of the (n_3+d)-vector. (n_3 is the order of $\mathscr{A}_3$ and d is the order of the error system). There is, for instance, the decision of whether to incorporate a given appendage mode into the model $\mathscr{A}_3$ and hence into the vector x^3, or whether to delete (truncate) the mode from $\mathscr{A}_3$ (x^3) and incorporate its effect in the "synthetic modes" associated with γ. The first difference to appreciate is that certain parameter uncertainties may be "forgiven" if the mode is accommodated within the synthetic modes of the error system as discussed in section 4.0. If all the parameters associated with the mode in question are quite well known and if the mode is known to have a significant effect upon the output then it may be desirable to leave the mode in $\mathscr{A}_3$. Usually in space-craft structures the modal data is somewhat in question. The frequencies of the assumed modes are often uncertain because of the approximations made in characterizing the model. For example, the frequencies of the forced appendage are not the same as those obtained by solving the eigenvalue problem for the homogeneous appendage equation as is often done. The structural damping of spacecraft, is another parameter which is difficult to predict reliably. A typical

135

pattern is to select a number for, ξ, the modal damping (after transformation to distributed coordinates and truncation is accomplished), in (6.2), and if possible perform some ground vibration tests. In such tests it is difficult and sometimes impossible to provide the boundary conditions indicative of orbital (free) flight, and the ground vibration data is suspect. In LST the task is further complicated by the unprecedented pointing requirements, and test data for damping with such small displacements will either be impractical to obtain or unreliable (or both). For these various reasons our first choice of an LST model $\mathcal{A}_3$ excludes all flexible modes so that they can be approximated by the synthetic modes of the error system. Also we wish to begin by considering the least complex controller (hence smallest model). Thus, the model $\mathcal{A}_3$ is taken to be, from (6.6) neglecting e,

$$\mathcal{A}_3 \begin{cases} x^3 = A^3 x^3 + B^3 u^o \\ y^3 = c^3 x^3 \\ z^3 = M^3 x^3 \end{cases} \tag{6.8}$$

where the pitch axis $\underline{b}_2$ is selected for study in this planar problem, and where

$$x^3 \triangleq \begin{pmatrix} \theta_2 \\ \dot{\theta}_2 \end{pmatrix} \qquad\qquad u^o = T_{c_2}$$

$$A^3 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad , \qquad B^3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$C^3 = M^3 = \qquad\qquad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

136

6.3 Solution of The Error Free Problem (STEP II)

The solution of the linear regulator problem for (6.8) follows.

$$\min_{u^o} \int_0^\infty (y^3 \, Q \, y^3 + u^{o^T} R \, u^o) dt$$

subject to

$$\dot{x}^3 = A^3 \, x^3 + B^3 \, u^o \quad , \quad A^3 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \, B^3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$
$$y^3 = c^3 \, x^3$$
$$z^3 = M^3 \, x^3 \qquad\qquad c^3 = I_2 \qquad\qquad (6.9)$$
$$x^3(o) = x_o$$

yields (since (A^3, c^3) is observable, (A^3, B^3) is controllable)

$$u^o = - R^{-1} B^{3^T} K \qquad\qquad (6.10)$$

where K obtains from

$$0 = - K A^3 - A^{3^T} K + K B^3 R^{-1} B^3 K - c^{3^T} Q c^3 \qquad\qquad (6.11)$$

The resulting K is

$$K = \begin{bmatrix} \sqrt{q_1(q_2 + 2\sqrt{q_1})} & \sqrt{q_1} \\ \sqrt{q_1} & \sqrt{q_2 + 2\sqrt{q_1}} \end{bmatrix} \qquad\qquad (6.12)$$

where Q and R have been selected, in form,

$$R = 1 \qquad\qquad (6.13)$$

$$Q = \begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix} \qquad\qquad (6.14)$$

Noting that the closed loop system for (6.9), (6.10) is

$$\dot{x}^3 = \bar{A}^3 \, x^3$$

where

137

$$\overline{A}^3 \begin{bmatrix} 0 & 1 \\ -\sqrt{q_1} & -\sqrt{q_2 + 2\sqrt{q_1}} \end{bmatrix} \tag{6.15}$$

We can relate the parameters q_1, q_2 to the classical parameters for a second order system, ξ, ω_n, by writing $\overline{A}^3$ in the form

$$\overline{A}^3 = \begin{bmatrix} 0 & 1 \\ -\omega_n^2 & -2\xi\omega_n \end{bmatrix} \tag{6.16}$$

and equating (6.15) to (6.16). The result is

$$\begin{aligned} q_1 &= \omega_n^4 \\ q_2 &= 2\omega_n^2 (2\zeta^2 - 1) \end{aligned} \tag{6.17}$$

Thus, if we wish, in this case, to choose the matrix Q on the basis of the classical parameters ζ and ω_n, we may do so since, from (6.14) (6.17)

$$Q = \begin{bmatrix} \omega_n^4 & 0 \\ 0 & 2\omega_n^2(2\zeta^2 - 1) \end{bmatrix} \tag{6.18}$$

Note further that the requirement for Q to be positive definite implies, from (6.18),

$$\zeta \geq .707 \tag{6.19}$$

for optimal solutions to (6.9). It is also helpful to rewrite (6.12) using (6.17)

$$K = \begin{bmatrix} 2\xi\omega_n^3 & \omega_n^2 \\ \omega_n^2 & 2\xi\omega_n \end{bmatrix} \tag{6.20}$$

Finally, the last required calculation of STEP II is

$$G_1 \triangleq - R^{-1} B^{3^T} K = (- \omega_n^2 \ , \ -2\xi \, \omega_n) \tag{6.21}$$

Rather than specify in (6.21) the free parameters ξ, ω_n (and thereby place the poles of $\mathscr{A}_3$) immediately we will (since the model is so simple) carry them in parametric form for a while.

In section 4.0 it was mentioned that deterministic designs by optimization have no natural band limiting quality. It is interesting to note in this simple example that the optimization solution (6.15) allows pole placement anywhere in the region R^3 of the complex plane
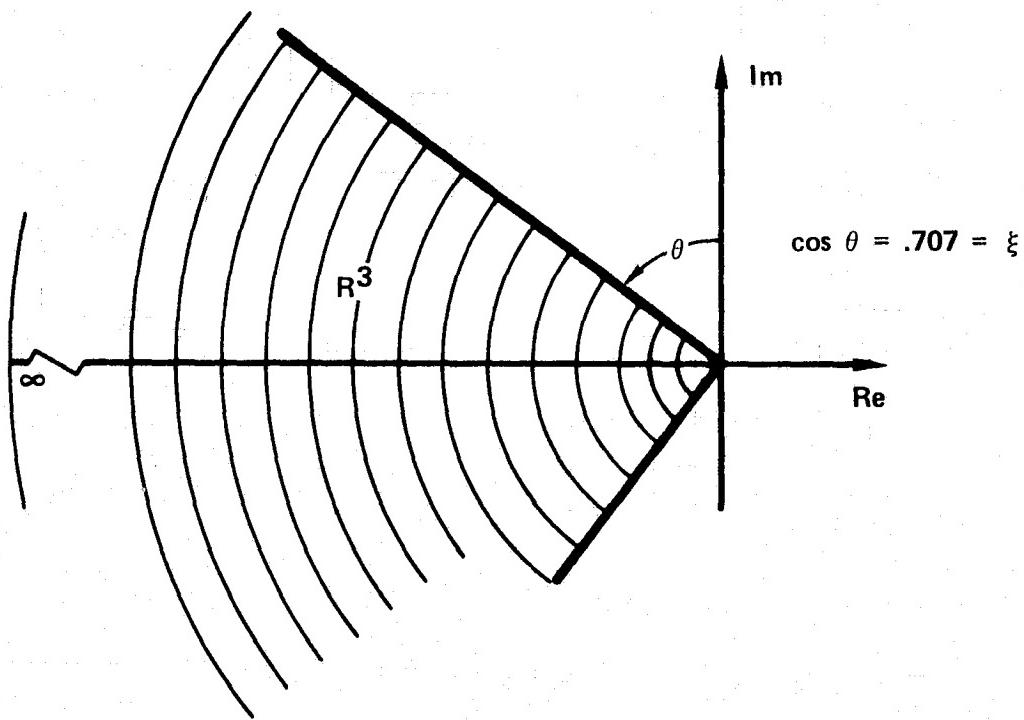
Quite obviously, however, the fidelity of the rigid body model $\mathscr{A}_3$ is questionable beyond the spectrum of truncated modes. That is, from Appendix C, the eigenvalue magnitude of the truncated rate gyro modes, (from (C-31)),

$$|\lambda_{rg}| = \omega_r = 2\pi \; 16$$

and the eigenvalue magnitude of the neglected position sensor mode (from (C-33)),

$$|\lambda_{gs}| = 2\pi$$

and the eigenvalue of the truncated A/D prefilter mode (from (C-35)),

$$|\lambda_p| = 2\pi \; 20$$

and the eigenvalue of the first appendage mode (estimated)

$$|\lambda_{FRUSA}| = 2\pi(.1)$$

all establish bounds on the spectrum of credibility, $C(\mathscr{A}_3)$, of the model $\mathscr{A}_3$; see Figure 6-2a.

The Kalman filter design for model $\mathscr{A}_3$ provides further insight into the effects of modeling errors. We will indulge in a brief digression now to see where the Kalman filter for the estimation of x^3 places the estimator poles relative to the above spectrum of credibility for model $\mathscr{A}_3$. Using the $\mathscr{A}_3$ parameters in (6.9) and the Kalman filter equations of Figure 2-2, it can be shown that the steady state (i.e., $\dot{\Sigma} = 0$ in Figure 2-2) Kalman filter,

$$\dot{\hat{x}}^3 = d \; \hat{x}^3 + B^3 \; u^o + \hat{K} \; z^o$$

$$d \; \underline{\underline{\Delta}} \; [A^3 - \hat{K} \; M^3] \tag{a}$$

has the eigenvalues

$$\lambda = \frac{1}{2} \; \mathrm{tr} \; d \left[1 \pm \sqrt{1 - \frac{4/d/}{(\mathrm{tr} \; d)^2}} \; \right] \tag{b}$$

where
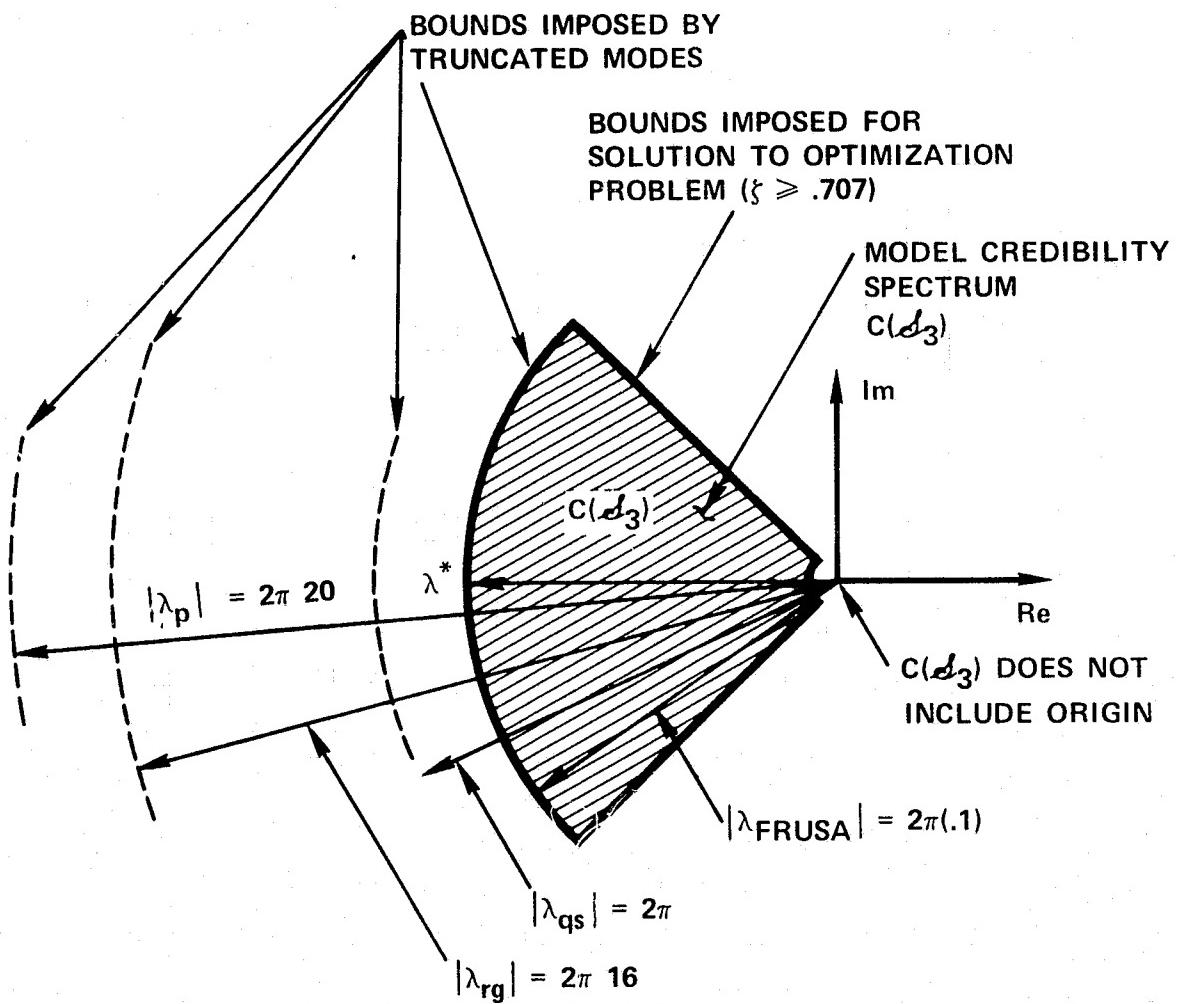
Figure 6-2a. LST Model Credibility Spectrum, $C(\mathcal{A}_3)$.

$$\text{tr } d = - \left(\frac{q}{r_1}\right)^{1/4} \left[\frac{\sqrt{qr_1}\,(qr_1 - 3r_2^2) - 2r_2^3}{(r_2 - \sqrt{q\,r_1})^2\,r_2}\right]^{1/2}$$

$$|d| \triangleq \det d = - \left(\frac{q}{r_1}\right)^{1/2}$$

(c)

where

r_1 = variance of the position sensor noise

r_2 = variance of the rate gyro noise

q = variance of the disturbance torque noise

$$\hat{R} \triangleq \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix} \quad, \quad \hat{Q} \triangleq q$$

Note that the estimator poles leave the model credibility spectrum $C(\mathcal{A}_3)$ as $q \to 0$ and again as $\lambda(qr) > \lambda^*$ (see Figure 6-2a). Under these conditions the Kalman filter might diverge, because it is making predictions over time intervals (too long or too short) for which the model is not reliable. To see this more clearly, consider that the right boundary of $C(\mathcal{A}_3)$ corresponds to time constants $\left.\frac{1}{\lambda}\right|_{\lambda=0} = \tau \to \infty$, Now since the filter relies absolutely upon ("believes") the model it has built within, it places[*] the estimator poles at the origin (corresponding to $q/r_1) \to 0$ in (c)). This action serves to disconnect the measurements from the filter (that is $\hat{K} = 0$ in Figure 2-2) since the filter feels justified in making predictions over arbitrarily long intervals, $\tau \to \infty$. Now when the estimator places its poles at (or too

***

[*]Recall that these events unfold in time according to the differential equation for $\hat{K}$ in Figure 2-2, evolving finally to the steady state values described above.

near) the origin, (to the right of $C(\mathcal{A}_3)$) it can diverge because of predictions over longer intervals than that for which the model is reliable. When the estimator places its poles to the left of $C(\mathcal{A}_3)$, the estimator may diverge again due to predictions over intervals inappropriate for the model (but in this event the intervals, τ_s, are too short, $1/\lambda = \tau_r$, $\lambda \geq \lambda$). In view of this problem we might naturally ask the following questions.

1. Can this estimator be stabilized by better statistical data for the noise? Not necessarily. Assume that q, r_1, r_2 <u>actually</u> represent the properties of the physical system. It might still happen that the pole placement specified by (a), (b), falls outside $C(\mathcal{A}_3)$ (note that information concerning the boundaries of $C(\mathcal{A}_3)$ is <u>not</u> contained in (a)).

2. Can this estimator be stabilized by better plant parameters? Not necessarily. Just as with the statistical parameters discussed above, there may not exist <u>any</u> plant parameters (A^3, M^3, B^3) which will stabilize the divergent Kalman filter (see Reference 22 for further examples of this circumstance).

3. Can this estimator be stabilized by the adaptive Kalman filter? Again, not necessarily. As discussed in section 3.0 most adaptive filtering schemes increase q in some way. Note that increasing q may help if the estimator poles are initially near the origin. However, if q, r_1, r_2 correspond to poles near the left boundary of $C(\mathcal{A}_3)$ the adaptive Kalman filter will actually cause the state estimates to get worse.

The conclusion drawn from this digression with the stochastic problem solution for $\mathcal{A}_3$ is that poor performance from a Kalman filter does <u>not</u> imply a poor choice of parameters for either the noise covariances or for the plant.* The Kalman filter ignores the presence of boundaries of $C(\mathcal{A}_3)$. The adaptive Kalman filter shares a fault with deterministic estimators of Luenberger that the presence of a left boundary of $C(\mathcal{A}_3)$ for this example, would be ignored. From this example we may conclude also that the deterministic (Luenberger) estimator does <u>not</u> have to be "faster" than the control system, as is typically suggested.[55,57] This follows from the result which shows the equivalence between the steady state Kalman estimator and the Luenberger estimator. From the note that the Kalman filter poles might fall to the left <u>or to the right</u> of the poles of the closed loop control system depending upon the selection of ξ, ω_n in (6.18) and the (independent) parameter values q, r_1, r_2 in (c).

6.4 The Characteristic Time of $\mathcal{A}_3$ (STEP III)

From the derivation of characteristic time in (4.78),

$$\tau = \lambda_{max} \{ [K \; B^3 R^{-1} B^3 K + C^{3^T} Q \; C^3]^{-1} K \} \tag{6.22}$$

$$= \lambda_{max} \{K\}$$

where from (6.9), (6.13), (6.18), and (6.20), the eigenvalues of K are

$$\lambda\{K\} = \frac{4\zeta^2 - 1}{2\omega_n(3\zeta^2 - 1)} \left[\zeta \pm \frac{2\zeta^2 - 1}{\sqrt{4\zeta^2 - 1}} \right]$$

---

*The more likely case is that the means of the random variables w, ν are not zero (nor or they necessarily constant). The method to model such correlated disturbances is to use the structure of the error system proposed in this study, specifically in section 5.0.

Then

$$\tau = \lambda_{max}\{K\} = \frac{4\zeta^2 - 1}{2\omega_n(3\zeta^2 - 1)} \left[\zeta + \frac{2\zeta^2 - 1}{\sqrt{4\zeta^2 - 1}} \right] \qquad (6.23)$$

Note that the eigenvalues of the closed loop $\mathscr{A}_3$ system are

$$\lambda[\overline{A}^3] = -\omega_n[\zeta \pm \sqrt{\zeta^2 - 1}]$$

6.5 Characterization of The Error Vector $e^{31}(t)$ (STEP IV)

In the absence of other information one may assume a "mode shape" for $e^{31}(t)$, as one alternative discussed in section 4.0. We choose, however, to calculate directly that e^{31} which is defined in (6.6), (6.7) (actually, $e^{31} \triangleq e_2$ in (6.6)). From the approximate data given in Table 6-1, (6.7) yields

$$e^{31}(t) = .3e^{-.05t} + .3e^{-.006t} \cos .628t$$

$$+ .007 \sin .00111 \, t - .08 \sin .00222 \, t + .007 \sin .00333t$$

$$+ .007 \sin .00444t \qquad (6.24)$$

where we have used the calculation of the external disturbances, T_o, in Newton meters,

$$T_o = -.034 \sin .00222 \, t + .02 \cos .00222 \, t$$

$$.003 \sin .00111 \, t + .003 \sin .00222 \, t +$$

$$.003 \sin .00333 \, t + .003 \sin .00444 \, t \qquad (6.25)$$

which is an approximation of (C-69) in Appendix C.

Table 6-1

LST PARAMETERS ASSUMED FOR SOLAR PANEL MODE

$\Gamma\Phi$ = 15.8

ζ = .01   modal damping

σ = $2\pi(.1)$ rad/sec   solar panel mode natural frequency

m_1 = 10.0 Kg   modal mass

F = J_2 = 46000 Kg m^2 $\underline{b}_2$ axis inertia

λ = .05 1/sec   thermal time constant

Δ_o = 1.0 m   steady state thermal deflection

Ω_o = .00111 rad/s   orbital rate

m = 8,000 Kg   vehicle mass

6.6 Selecting The Synthetic Modes (STEP V)

Here we choose the Chebyshev error system

$$\hat{e}^{31} = P\gamma$$
$$\dot{\gamma} = D\gamma$$

where

$$D = \frac{2}{\tau} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 3 & 0 & 6 & 0 \\ & & & & \ddots \end{bmatrix} dd \qquad (6.26)$$

and P is defined next.

6.7 The Chebyshev-Fourier Coefficients, $P(\tau, d)$ (STEP VI)

From (5.12)

$$P = \int_{-1}^{1} g(\sigma) \ e^{31}(\sigma) \ T^T(\sigma) d\sigma \ \Lambda^{-1} \qquad (6.27)$$

$$\Lambda^{-1} = \frac{2}{\pi} \begin{bmatrix} 1/2 & 0 \\ 0 & I_{d-1} \end{bmatrix}$$

Using the Chebyshev quadrature formula in (A-17) for evaluating (6.27) we obtain, using the data in Table 6-1, and $e^{31}(t)$ given by (6.24),

$$P = \begin{bmatrix} 0 & 0 & 0 & 0 \\ P_1 & P_2 & P_3 & P_4 \end{bmatrix}_{2d} \quad \ldots \ldots \qquad (6.28)$$

where

$$p_1 = .645$$

$$p_2 = -.155$$

$$p_3 = .62$$

$$p_4 = .05$$

where we have now selected for ξ, ω_n in (6.17)

$$\zeta = 1.0$$
$$\omega_n = .4 \qquad (6.29)$$

6.8 Feedback Gains for the Synthetic Variables of the Error System (STEP VII)

Now we solve for L from (5.13).

$$\left[K \ B^3 R^{-1} B^{3^T} - A^{3^T} \right] L - LD = KP \qquad (6.30)$$

The recursive solution to (6.30) given by the lemma of section 5.1 yields (define $S^3 \triangleq B^3 R^{-1} B^{3^T}$)

$$L^d = \left[KS^3 - A^{3^T}\right]^{-1} K P^d \qquad (6.31)$$

$$L^{d-1} = \left[KS^3 - A^{3^T}\right]^{-1} \left(KP^{d-1} + L^d D_{d(d-1)}\right)$$

$$L^{d-2} = \left[KS^3 - A^{3^T}\right]^{-1} \left(KP^{d-2} + L^{d-1} D_{(d-1)(d-2)}\right)$$

$$L^{d-3} = \left[KS^3 - A^{3^T}\right]^{-1} \left(KP^{d-3} + L^{d-2} D_{(d-2)(d-3)} + L^d D_{d(d-3)}\right)$$

$$L^{d-4} = \left[KS^3 - A^{3^T}\right]^{-1} \left(KP^{d-4} + L^{d-3} D_{(d-3)(d-4)} + L^{d-1} D_{(d-1)(d-4)}\right)$$

$$\vdots$$

Now since

$$\left[K^3 S^3 - A^{3^T}\right]^{-1} = \begin{bmatrix} \dfrac{2\xi}{\omega_n} & -1 \\[2ex] \dfrac{1}{\omega_n^2} & 0 \end{bmatrix} \qquad (6.32)$$

and

$$\left[K^3 S^3 - A^{3^T}\right]^{-1} K^3 = \begin{bmatrix} (4\zeta^2 - 1)\omega_n^2 & 0 \\[2ex] 2\zeta\,\omega_n & 1 \end{bmatrix} \qquad (6.33)$$

then the columns of L become, noting the values of D_{ij} in (5.11),

<u>for d = 1</u>, (denote L by L_1)

$$L_1^1 L^1 = \begin{pmatrix} 0 \\ P_1 \end{pmatrix} \qquad (6.34)$$

For d = 2, (denote L by L_2)

148

$$L_2^2 = \begin{pmatrix} 0 \\ p_2 \end{pmatrix} \tag{6.35}$$

$$L_2^1 = \begin{pmatrix} 0 \\ p_1 \end{pmatrix} + \begin{pmatrix} -p_2 \\ 0 \end{pmatrix} \frac{2}{\tau} = \begin{pmatrix} -\dfrac{2}{\tau}\, p_2 \\ p_1 \end{pmatrix} \tag{6.36}$$

For $d = 3$, (denote L by L_3)

$$\left. \begin{aligned}
L_3^3 &\quad \begin{pmatrix} 0 \\ p_3 \end{pmatrix} \\[2mm]
L_3^2 &= \begin{pmatrix} 0 \\ p_2 \end{pmatrix} + \begin{pmatrix} -p_3 \\ 0 \end{pmatrix} 4\,\frac{2}{\tau} = L_2^2 + \begin{pmatrix} -p_3 \\ 0 \end{pmatrix} \frac{8}{\tau} \\[2mm]
L_3^1 &= \begin{pmatrix} 0 \\ p_1 \end{pmatrix} + \begin{pmatrix} -p_2 \\ 0 \end{pmatrix} \frac{2}{\tau} + \begin{pmatrix} -\dfrac{2\zeta}{\omega_n} \\ -\dfrac{1}{\omega_n^2} \end{pmatrix} p_3\, 4 \left(\frac{2}{\tau}\right)^2
\end{aligned} \right\} \tag{6.37}$$

$$L_3^1 = L_2^1 - \frac{16}{\tau^2 \omega_n^2}\, p_3 \begin{pmatrix} 2\zeta\,\omega_n \\ 1 \end{pmatrix}$$

Thus, as d is increased, the old values of L are retained and a correction term is added.

For the matrix G we must compute

$$G = \left[-R^{-1}B^{3^T}K^3 \;,\; -R^{-1}B^{3^T}L \right] = \left[G_1, \; G_2 \right] \tag{6.38}$$

Hence, from the data (6.1), (6.3) and $R = 1$,

$$-R^{-1}B^{3^T}K^3 = -\left(\omega_n^2 \;,\; 2\zeta\,\omega_n \right) \tag{6.39}$$

$$-R^{-1}B^{3^T}L = -\left(L_{21}, \; L_{22}, \ldots, L_{2d} \right) \tag{6.40}$$

From (6.12), (6.14), (6.15), (6.13) can be written, for $d \leq 3$,

$$G = \left(-\omega_n^2, \ -2\zeta\,\omega_n, \ -P_{21} + \frac{16}{\tau^2 \omega_n^2} P_{23} \, , \right.$$

$$\left. -P_{22} \, , \ -P_{23} \, , \cdots \cdot \right) \tag{6.41}$$

and

$$G_2 = \left(-P_1 + \frac{16}{\tau^2 \omega_n^2} \ P_3 \, , \ -P_2 \, , \ -P_3 \, , \cdots \right) \tag{6.42}$$

where $p_i = 0$ for $i > d$.

6.9 Estimator Gain Determination (STEP VIII)

In order to gain the insight available from an analytical computation of the estimator gains, choose $d = 2$, a tractable case. Now first we will follow a simplified version of STEPS VIIIa, IXa, listed in section 5.2.

To actually get a meaningful evaluation of the controller design a more substantial "evaluation" model should be used than the model $\mathcal{A}_1$ used here. We will here simply show a choice of estimator gains which makes (5.16) a stability matrix when j=3. The alternate STEP VIIIb may also be selected for the task of stabilizing the estimator (with perhaps even less effort).

Now for the matrix A in (5.16), using the model $\mathcal{A}_3$ given in (6.9)

$$A = \left[\begin{array}{c|c} A^3 + B^3(G_1 + G_2 \, T_{21}) & B^3 \, G_2 \\ \hline DT_{21} - T_{21} \, (A^3 + B^3 G_1) & D \end{array} \right] \tag{6.43}$$

Define the parameters within T_{21}

$$T_{21} = \begin{bmatrix} 0 & t_2 \\ 0 & t_4 \end{bmatrix} \tag{6.44}$$

where the zero entries result from the fact that the matrix P has a zero first row (hence the eigenvalues of the estimator, $\lambda[d_{22}] = \lambda[D - T_{21}P]$, are independent of the first column of T_{21}). The closed loop system is now described by two physical variables θ_2, $\dot{\theta}_2$, and two synthetic variables of the error system state estimates $\hat{\gamma}_1$, $\hat{\gamma}_2$. This system is written, from (4.115), using (6.42)-(6.44),

$$\begin{pmatrix} \dot{\theta}_2 \\ \ddot{\theta}_2 \\ \dot{\hat{\gamma}}_1 \\ \dot{\hat{\gamma}}_2 \end{pmatrix} = \left[\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ -\omega_n^2 & -2\xi\omega_n - p_1 t_2 - p_2 t_4 & -p_1 & -p_2 \\ \hline \omega_n^2 t_2 & 2\zeta\omega_n t_2 & 0 & 0 \\ \omega_n^2 t_4 & 2\zeta\omega_n t_4 + \frac{2}{\tau} t_2 & \frac{2}{\tau} & 0 \end{array} \right] \begin{pmatrix} \theta_2 \\ \dot{\theta}_2 \\ \hat{\gamma}_1 \\ \hat{\gamma}_2 \end{pmatrix} +$$

$$\begin{bmatrix} 0 & 0 \\ -\omega_n^2 & -2\zeta\omega_n - p_1 t_2 - p_2 t_4 \\ \omega_n^2 t_2 & 2\zeta\omega_n t_2 \\ \omega_n^2 t_4 & 2\zeta\omega_n t_4 + \frac{2}{\tau} t_2 \end{bmatrix} \begin{pmatrix} \tilde{\theta} \\ \dot{\tilde{\theta}} \end{pmatrix} + \begin{pmatrix} 0 \\ e(t) \\ 0 \\ 0 \end{pmatrix} \tag{6.45}$$

where $\tilde{\theta}$, $\dot{\tilde{\theta}}$ represents, as did $\tilde{z}^3$ in (4.115), the measurement residual errors (the difference between the actual measurements and those predicted by the model). As noted in prior discussions these residuals may be characterized as noise in the estimator gain evaluation, if desired. The transfer function between the position $\theta_2(s)$ and the model error, $R(s)$, is, from (6.45),

$$\frac{\theta_2(s)}{e(s)} = \frac{s^2}{s^4 + \alpha_3 s^3 + \alpha_2 s^2 + \alpha_1 s + \alpha_o} \qquad (6.46)$$

where

$$\alpha_3 \triangleq 2\zeta\,\omega_n + 2\zeta_e\omega_e$$

$$\alpha_2 \triangleq \omega_n^2 + 2\zeta\,\omega_n\,(2\zeta_e\omega_e)$$

$$\alpha_1 \triangleq \omega_n^2\,2\zeta_e\omega_e + 2\zeta\,\omega_n\,\omega_e^2$$

$$\alpha_0 \triangleq \omega_n^2\,\omega_e^2$$

and where we have substituted in (6.46) the more convenient parameters

$$\omega_e^2 \triangleq \frac{2}{\tau}\,t_2\,p_2$$

$$2\zeta_e\omega_e \triangleq t_2\,p_1 + t_4\,p_2 \qquad (6.47)$$

Furthermore, the eigenvalues of the estimator, which are derived from

the matrix d_{22} (see (4.91)), may be expressed in terms of ζ_e, ω_e

$$\lambda[d_{22}] = \omega_e\left[-\zeta_e \pm \sqrt{\zeta_e^2 - 1}\,\right] \qquad (6.48)$$

A Routh analysis of (6.46) indicates that for stability of (6.45)

there is a lower limit for the "speed" of the estimator, ω_e, relative

to the "speed" of the controlled error free system, ω_n,

$$r \triangleq \frac{\omega_e}{\omega_n} > \frac{2\zeta\,\zeta_e}{4\xi_e^2 - 1}\left[-1 + \sqrt{1 - \frac{4\,\xi_e^2 - 1}{4\,\zeta^2\,\zeta_e^2}}\,\right] \qquad (6.49)$$

and, consistent with our expectations, there is no upper limit for ω_e,

(based upon these stability arguments).

Now the controller which must be implemented according to this

design is illustrated in Figure 6-2 in block diagram form and given
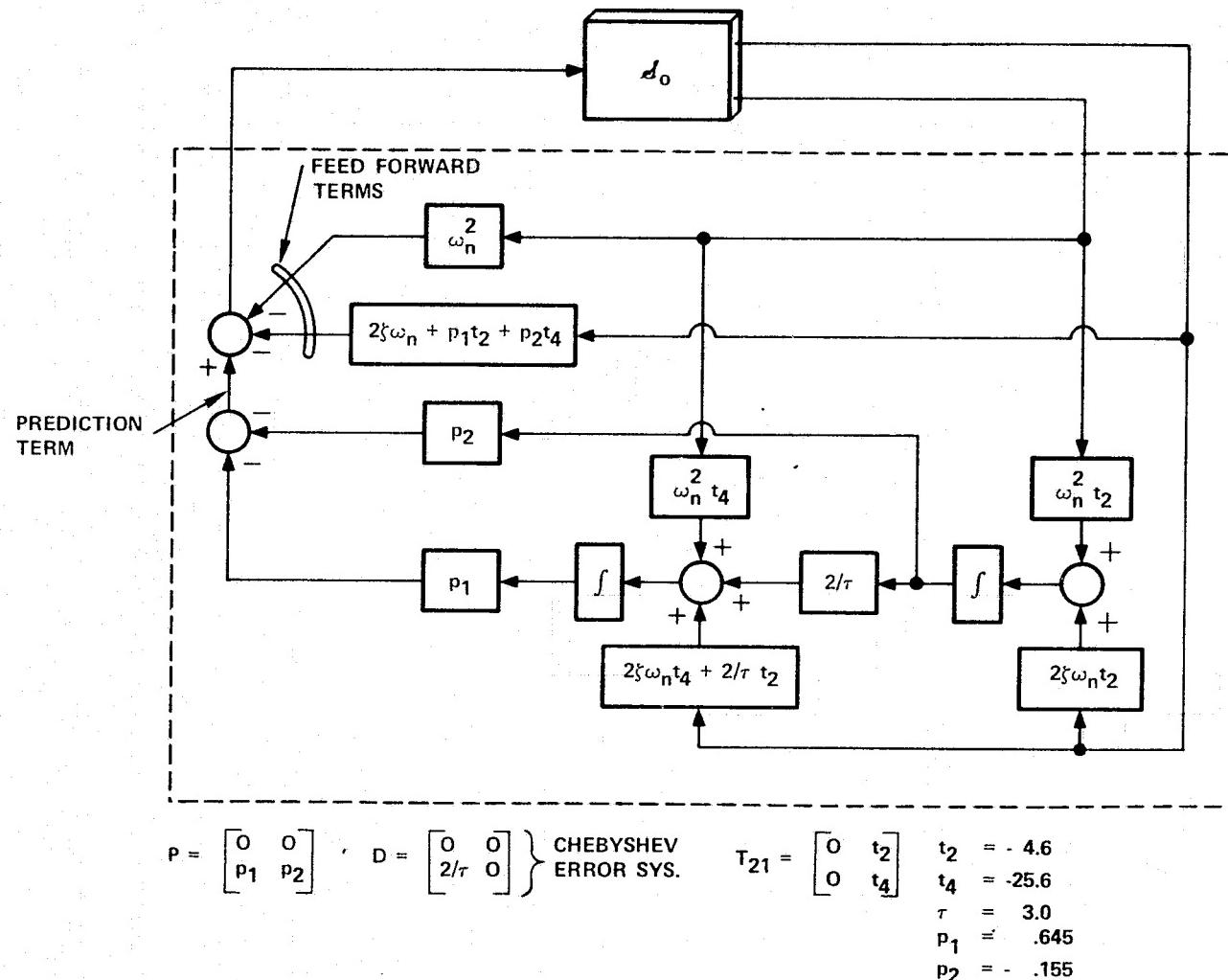
Figure 6-2. Orthogonal Filter for LST Using Chebyshev Polynomials.

below in equation form, from (4.111),

$$\begin{pmatrix} \dot{\hat{\gamma}}_1 \\ \dot{\hat{\gamma}}_2 \end{pmatrix} = \begin{bmatrix} 0 & 0 \\ 2/\tau & 0 \end{bmatrix} \begin{pmatrix} \hat{\gamma}_1 \\ \hat{\gamma}_2 \end{pmatrix} + \begin{bmatrix} \omega_n^2 \, t_2 & 2\zeta \, \omega_n \, t_2 \\ \omega_n^2 \, t_4 & 2\zeta \, \omega_n \, t_4 + \frac{2}{\tau} \, t_2 \end{bmatrix} \begin{pmatrix} \theta_2^0 \\ \dot{\theta}_2^0 \end{pmatrix}$$

$$u^0 = (-p_1, -p_2) \begin{pmatrix} \hat{\gamma}_1 \\ \hat{\gamma}_2 \end{pmatrix} + \left(-\omega_n^2, \; -2\zeta \, \omega_n - t_2 p_1 - t_4 p_2 \right) \begin{pmatrix} \theta_2^0 \\ \dot{\theta}_2^0 \end{pmatrix}$$

(6.50)

The state estimator, (4.91) (and the lower part of (6.45)) and the closed loop evaluation model, (6.45)(from (4.115)), are both stabilized by the choices

$$\left. \begin{array}{l} \xi = 1.0 \\ \omega_n = .4 \end{array} \right\} \Rightarrow \tau = 3.0$$

$$t_2 = -4.6$$

$$t_4 = -25.6$$

It is further noted from (6.46) that due to the s^2 term in the numerator the vehicle pointing position will be insensitive to any model errors, $e(t)$, which are closely approximated (within time constants of (6.46)) by a constant or a ramp in time. Now, decreasing $(\omega_n \, \omega_e)$ will give better noise rejection, but for the larger time constants associated with smaller $\omega_n \, \omega_e$, fewer model errors can be approximated by a constant and a ramp. Consideration of sensor noise data, when available for substitution into (6.45) (for $\tilde{\theta}, \dot{\tilde{\theta}}$) will permit perhaps better choices for ζ, ω_n, t_2, t_4 from the multi-model procedure described above. The competition between the "low frequency" and the "high frequency" effects of modeling errors indicates that a "best" set of design parameters exists. Generally, however, neither

the Kalman filter nor the Luenberger observer designs provide these best parameters, owing to their commitment to a fixed model whose resulting controller performance is eventually limited by the errors of the model (such as truncated modes).

7.0 Concluding Remarks

This research has been concerned with that deficiency of modern control theory which causes it to fail to recognize limits to the fidelity of the mathematical model. A modeling plan is proposed for model construction which allows a parameter of the control problem (called the characteristic time, τ) to influence the modeling problem. This is accomplished through the introduction of an error system which is augmented to the model of the system and which is designed to approximate the effects of model errors (external disturbances, truncated modes and parameter errors) only over the interval, τ. The error system chosen belongs to a class called the Sturm Liouville equations which are known to generate orthogonal functions. The benefit such an error system provides to the closed loop linear regulator controller (using linear state estimation for feedback of the states of the model and the states of the error system) is that when measurements from the physical system differ from those predicted by the model (built within the state estimator) the difference is automatically fitted with a set of orthogonal functions (for a least squares fit over the interval τ) and the corrected signal is used for control. The estimator for the error system is therefore called an "orthogonal filter" and is used for control by model error estimation.

This report lends further support to prior claims[1-6] that increasing the order of the model can provide better systems performance. When the "model error vector" described above is considered to result only from external disturbances, then these procedures can be used to design "disturbance accommodating controllers," discussed in

PRECEDING PAGE BLANK NOT

Reference 4 when the "model error vector" is considered simply as white
noise the Kalman filter can be obtained. The method has a computa-
tional requirement similar to that of the singular perturbation
approach,[8] in that a Riccati equation need be solved only in the
smaller dimension of the "error free" system model (corresponding to
the "truncated" model of Reference 8).

An example illustrates the application of orthogonal filters in
the special case when the orthogonal approximating functions are
selected to be Chebyshev polynomials. The Chebyshev error system
that results is then applied to the control of the Large Space Tele-
scope (LST) which the National Aeronautics and Space Administration
(NASA) plans to use for extragalactic research in the 1980's.

8.0 Future Research

The following areas seem fruitful for further research.

1. Model Learning Observers: The model error vector, $e^{31}(t)$, was defined to be that vector which makes the output of the smaller model equal to the output of a larger, more credible model. The approximations used for $e^{31}(t)$, which were called $\hat{e}^{31}(t)$, do not in fact render the two outputs exactly equal. The measurement residual which results could be further approximated by considering the model error vector to influence the measurements directly through some matrix M_γ

$$z^2 = M^3 x^3 + M_\gamma \gamma \ , \ (M_\gamma \gamma \approx \tilde{z}^2)$$

which matrix, M_γ, could be adaptively regulated to decrease the measurement residual, $z^0(t) - z^2(t)$. Such a procedure, if found, would make the estimator "model learning" in the sense that better knowledge would be continuously obtained about how the model error vector actually influences the measurements.

2. In the Tri-Model design approach no simple procedure has been established for the solution of the estimator gains, T_{21}, from

$$\min_{T_{21}} J(T_{21}) = \min_{T_{21}} \int_0^T y^{1^T} Q y^1 \, dt$$

when the output y^1 evolves from a different model, $\mathscr{A}_1$, than that used for the estimator design, $\mathscr{A}_2$. When $T \to \infty$ difficulties can arise in evaluating $J_{(T_{21})}$ (see above) when the model error system is unstable. (Recall according to an evaluation of the closed loop system with

Model $\mathscr{A}_2$, the eigenvalues are equal to those of the error system in addition to those of the optimally controlled error free system). From a practical engineering point of view such solutions might well exist, and further work is needed to provide practical approximations for T_{21} even though $J^*(T_{21}) = \min_{T_{21}} \int_0^\infty y^{1^T} Q y^1 \, dt$ might not exist mathematically.

3. Further reductions in the order of the controller can sometimes be realized by estimating the linear functional $G\hat{x}^2$ instead of first estimating x^2 and then multiplying by G. Possible reductions in controller order should be investigated for the special matrix structure of the orthogonal filters.

4. Other practical problems should be solved using Chebyshev and Fourier error systems to determine some guidelines to indicate which orthogonal functions one should select. The limitations of the error system approach should be investigated. It seems likely that stability considerations will eventually limit the order of the error system (and equivalently the degree of the orthogonal functions) which can be augmented to the system model.

CITED REFERENCES

1. Astrom, K.J. and P. Eykhoff. "System Identification - A Survey," Automatica, Vol. 7, pp. 123-162, 1971.

2. Johnson, C.D. "Further Study of the Linear Regulator with Disturbances - The Case of Vector Disturbances Satisfying a Linear Differential Equation," IEEE Transactions on Automatic Control, Vol. AC - 15, No. 2, pp. 222-228, April 1970.

3. Johnson, C.D. and R.E. Skelton. "Optimal Desaturation of Momentum Exchange Control Systems," Proceedings of Joint Automatic Control Conference, Atlanta, Georgia, pp. 683-694, June 1970.

4. Johnson, C.D. "Accommodation of External Disturbances in Linear Regulator and Servomechanism Problems," IEEE Transactions on Automatic Control, Vol. AC-16, No. 6, pp. 635-645, December 1971.

5. Pearson, J.B. "Compensator Design for Dynamic Optimization," International Journal of Control, Vol. 9, No. 4, pp. 473-482, 1969.

6. Davison, E.J. "The Output Control of Linear Time-Invariant Multivariable Systems with Unmeasurable Arbitrary Disturbances," IEEE Transactions on Automatic Control, Vol. AC-17, No. 5, pp. 621-630, October 1972.

7. Kokotovic, P.V. and P. Sannuti. "Singular Perturbation Method for Reducing the Model Order in Optimal Control Design," IEEE Transactions on Automatic Control, Vol. AC-13, No. 4, pp. 377-384, August 1968.

8. Kokotovic, P.V. and R.A. Yackel. "Singular Perturbation of Linear Regulators: Basic Theorems," IEEE Transactions on Automatic Control, Vol. AC-17, No. 1 , pp. 29-37, February 1972.

9. Sannuti, P. and P.V. Kokotovic. "Near-Optimum Design of Linear Systems by a Singular Perturbation Method," IEEE Transactions on Automatic Control, Vol. AC-14, No. 1, pp. 15-22, February 1969.

10. Davison, E.J. "A New Method for Simplifying Large Linear Dynamic Systems," IEEE Transactions on Automatic Control, Vol. AC-13, No. 2, pp. 214, April 1968.

11. Rogers, R.O. and D.D. Sworder. "Suboptimal Control of Linear Systems Derived from Models of Lower Dimension," AIAA Journal, Vol. 9, No. 8, pp. 1461-1467, August 1971.

12. Sinha, N.K. and H. deBruin. "Near-Optimal Control of Higher Order Systems Using Low-Order Models," International Journal of Control, Vol. 6, No. 2, pp. 257-262, February 1973.

CITED REFERENCES (Continued)

13. Wonham, W.M. "On Pole-Assignment in Multi-Input Controllable Linear Systems," *IEEE Transactions on Automatic Control*, Vol. AC-12, No. 6, pp. 660-665, December 1967.

14. Ho and Lee. "Identification of Linear Dynamical Systems," *Information and Control*, Vol. 8, pp. 93-110, 1965.

15. Kreindler and Sarachik. "On the Concepts of Controllability and Observability of Linear Systems," *IEEE Transactions on Automatic Control*, Vol. AC-15, No. 17, pp. 129-136, April 1964.

16. Shultz and Melsa. *State Functions and Linear Control Systems*, McGraw-Hill, New York, 1967.

17. Quigley, A.L.C. "An Approach to the Control of Divergence in Kalman Filter Algorithms," *International Journal on Control*, Vol. 17, No. 4, pp. 741-746, 1973.

18. Jazwinski, A.H. *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.

19. Wolf, H. "An Assessment of the Work in Minimum Variance Orbit Determination," Analytical Mechanics Associates, Inc., Westbury, New York, Report No. 68-3, January 1968.

20. Pines, et al. "Modification of the Goddard Minimum Variance Program for the Processing of Real Data," Analytical Mechanics Associates, Inc., Westbury, New York, Report Contract Nas 5-2535, October 1964.

21. Price, C.F. "An Analysis of the Divergence Problem in the Kalman Filter," *IEEE Transaction on Automatic Control*, Vol. AC-13, No. 2, pp. 699-702, December 1968.

22. Fitzgerald, R.J. "Divergence of the Kalman Filter," *IEEE Transactional on Automatic Control*, Vol. AC-16, No. 6, pp. 736-747, December 1971.

23. Griffin, R.E. and A.P. Sage. "Large and Small Scale Sensitivity Analysis of Optimum Algorithms," *IEEE Transactions on Automatic Control*, Vol. AC-13, No. 4, pp. 320-329, August 1968.

24. Fujita, S. and T. Fukao. "Error Analysis of Optimal Filtering Algorithm for Colored Noise," *IEEE Transactions on Automatic Control*, Vol. AC-15, No. 4, p. 452, August 1970.

25. Leondes, C.T. and J.O. Pearson. "A Minimax Filter for Systems with Large Plant Uncertainties," *IEEE Transactions on Automatic Control*, pp. 266-268, April 1972.

26. Guardabassi, G., A. Locatelli, and S. Rinaldi. "On the Optimality of the Wilkie-Perkins Low-Order Sensitivity Model," *IEEE Transactions on Automatic Control*, Vol. AC-15, pp. 382-384, June 1970.

27. Hontoir, Y. and J.B. Cruz, Jr. "Sensitivity Reduction in Linear Systems," Automatica, Vol. 8, No. 4, pp. 445-449, 1972.

28. Guardabassi, G., A. Locatelli, and S. Rinaldi. "Structural Properties of Sensitivity Systems," *Journal of the Franklin Institute*, Vol. 294, No. 4, pp. 241-249, October 1972.

29. Sundararajan, N. and J.B. Cruz, Jr. "Trajectory Insensitivity of Optimal Feedback Systems," *IEEE Transactions on Automatic Control*, Vol. AC-15, No. 6 , pp. 663-665, December 1970 (Short papers).

30. Guardabassi, G., A. Locatelli, and S. Rinaldi, "Further Results On Controllability and Sensitivity," *IEEE Transactions on Automatic Control*, Vol. AC-15, No.   , pp. 697-698, December 1970 (Correspondence).

31. Kang, H.S. "Minimum Sensitivity Design of M.V. Systems," *International Journal of Control*, Vol. 11, No. 5, pp. 791-801, 1970.

32. Newmann, M.M. "On Attempts to Reduce the Sensitivity of the Optimal Linear Regulator to a Parameter Change," *International Journal of Control*, Vol. 11, No. 6, pp. 1079-1084, 1970.

33. Guardabassi, G., A. Locatelli, and S. Rinaldi. "Structural Uncontrollability of Sensitivity System," *Automatica*, Vol. 5, pp. 297-301, May 1969.

34. Sakharov, M.P. "Quantitative Comparison of Types of Incompleteness of Sensitivity Models," *Automatic and Remote Control*, Vol. 33, No. 1, Pt. 1, pp. 79-86, January 1972.

35. Nayfeh, A.H. *Perturbation Methods*, John Wiley & Sons, New York, 1973.

36. Bhattacharyya, S.P. and J.B. Pearson. "On Error Systems and The Servomechanism Problem," *International Journal of Control*, 1972, Vol. 15, No. 6, pp. 1041-1062.

37. Athan, M. and Falb, P.L. *Optimal Control, An Introduction to the Theory and its Applications*, McGraw-Hill, New York, 1966.

38. Meditch, J.S., *Stochastic Optimal Linear Estimation and Control*, McGraw-Hill, New York, 1969.

CITED REFERENCES (Continued)

39. Special Issue on Linear-Quadratic-Gaussian Problem, _IEEE Transactions on Automatic Control_, Vol. AC-16, No. 6, December 1971.

40. Brockett, Roger W., _Finite Dimensional Linear Systems_, John Wiley New York, 1970.

41. Leondes, C.T. and J.O. Pearson, "Kalman Filtering of Systems with Parameter Uncertainties," _International Journal of Control_, Vol. 17, No. 4, pp. 785-801, 1973.

42. D'Appolito, J.A., Hutchinson, C.E., "A Minimax Approach to the Design of Low Sensitivity State Estimators," _Automatica_, Vol. 8, pp. 599-608, September 1972.

43. Griffin, R.E., Sage, A.P., "Sensitivity Analysis of Discrete Filtering and Smoothing Algorithms," _AIAA Journal_, Vol. 7, p. 1890, 1967.

44. Fagin, F.L., "Recursive Linear Regression Theory Optimal Filter Theory, and Error Analysis of Optimal Systems," _IEEE Intern. Conv. Record 12_, pp. 216-240, 1964.

45. Schmidt, S.F., "Estimation of State with Acceptable Accuracy Constraints," Analytical Mechanics Associates, Inc., Westbury, New York, Interin Report, 67-4, 1967.

46. Jazwinski, A.H., "Limited Memory Optimal Filtering," _IEEE Transactions on Automatic Control_, Vol. AC-13, No. 5, pp. 558-563, October 1968.

47. Bryson and Hendrikson, "Estimation Using Sampled-Data Containing Sequentially Correlated Noise," Tech. Rept. No. 533, Division of Engineering and Applied Physics, Harvard University, Cambridge, Mass., June 1967.

48. Stear, E.B., Stubberud, A.R., "Optimal Filtering for Gauss-Markov Noise," _International Journal of Control_, Vol. 8, No. 2, pp. 123-130, 1968.

49. Bhattacharyya, S.P. and J.B. Pearson. "On Error Systems and The Servomechanism Problem," _International Journal of Control_, Vol. 15, No. 6, pp. 1041-1062.

50. Moore, B.C., Silverman, L.M., "Model Matching by State Feedback and Dynamic Compensation," _IEEE Trans. Automatic Control_, Vol. AC-17, No. 4, p. 491, August 1972.

51. Chen, Yu, _Vibrations: Theoretical Methods_, Addison-Wesley.

CITED REFERENCES (Continued)

52. Anderson, B.D.D., and J.B. Moore, Linear Optimal Control, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1971.

53. Kalman, R.E. and Bertram, J.E. "Control System Analysis and Design Via the "Second Method" of Lyapunov, I Continuous-Time Systems," J. Basic Engr., pp. 371-393, June 1960.

54. Massera, J.L. "Contributions to Stability Theory," Ann. Math., Vol. 64, pp. 182-206, 1956.

55. Luenberger, D.G. "An Introduction to Observers," IEEE Transactions on Automatic Control, Vol. AC-16, No. 6, pp. 596-603, December 1971.

56. Kwatny, H.G. "Minimal Order Observers and Certain Singular Problems of Optimal Estimation and Control," IEEE Transactions on Automatic Control, Vol. AC-19, No. 3, pp. 274-276, June 1974.

57. Yuksel, Y.O., and J.J. Bongiorno, Jr., "Observers for Linear Multivariable Systems with Applications," IEEE Transactions on Automatic Control, Vol. AC-16, No. 6, pp. 603-613, December 1971.

58. Wonham, W.M. "On Pole Assignment in Multi-Input Controllable Linear Systems," IEEE Transactions on Automatic Control, Vol. AC-12, No. 6, pp. 660-666, December 1967.

59. Wolovich, W.A. "On The State Estimation of Observable Systems," in Proceedings of 9th Joint Automatic Control Conference, pp. 210-222, 1968.

60. Fortmann, T.E. and D. Williamson, "Design of Low-Order Observers for Linear Feedback Control Laws," IEEE Transactions on Automatic Control, Vol. AC-17, No. 7, pp. 301-308, June 1972.

61. K. Brammer, "Lower Order Optimal Linear Filtering of Nonstationary Random Sequences," IEEE Transactions on Automatic Control, Vol. AC-13, pp. 198-199, April 1968.

62. Novak, L.M. and C.T. Leondes, "Reduced Order Observers for Linear Discrete-Time Systems," IEEE Transactions on Automatic Control, Vol. AC-19, pp. 42-46, February 1974.

63. Szego, G. Orthogonal Polynomials, American Mathematical Society, 531 West 116th Street, New York, N.Y.

64. Ghizzetti, A. and A. Ossicini, Quadrature Formulae, Academic Press, pp. 124-125, New York 1970.

65. Krylov, V.I., Approximate Calculations of Integrals, MacMillan, New York, 1962.

CITED REFERENCES (Continued)

66. Abramowitz, M. and I.A. Stegun, Handbook of Mathematical Functions, National Bureau of Standards, Applied Mathematics Series 55, June 1966.

67. Likins, P.W. "Dynamics and Control of Flexible Space Vehicles," Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, Technical Report 32-1329, pp. 16-39, January 1970.

68. Skelton, R.C., Optimal Momentum Management in Momentum Exchange Control Systems for Orbiting Vehicles, MS in Engineering, Electrical Engineering Department, University of Alabama, Huntsville, Huntsville, Alabama, June 1969.

69. Luke, R.K.C. "Descriptions of the Geomagnetic Field," Aerospace Corporation, Los Angeles, California, Air Force Report No. SSD-TR-67-32, Aerospace Report No. TR-1001 (2307)-5, January 1967.

APPENDIX A   LEAST SQUARES FITTING OF DATA FOR MODEL ERROR

APPROXIMATION WITH JACOBI POLYNOMIALS

In this appendix we assume that the model error vector, $e(t)$, is a given function of time and we wish to approximate $e(t)$ by $\hat{e}(t)$ where

$$\hat{e} = Pf \qquad f^T = (f_1, \ldots, f_d)$$

and $f_i(t)$ are Jacobi polynomials. This class of polynomials is one of interest because as special cases within this class certain other desirable polynomials are obtained, such as the Fourier series, the Legendre and Chebyshev polynomials. Before specializing to these cases, however, we will first develop a model error system

$$
\begin{aligned}
e &= P\gamma \\
\dot{\gamma} &= D\gamma
\end{aligned}
\tag{A.1}
$$

which is capable of generating the more general Jacobi polynomials. It will be convenient in this Appendix to normalize time to vary over the interval $[-1, 1]$ instead of the interval $[0, \tau]$. Such a change of variable is accomplished by the definition

$$\sigma \triangleq 2\frac{t}{\tau} - 1 \qquad \begin{aligned} t &\in [0, \tau] \\ \sigma &\in [-1,1] \end{aligned} \tag{A.2}$$

The matrix P which minimizes

$$J(P) = \int_{-1}^{1} (e(\sigma) - \hat{e}(\sigma))^T q(\sigma) (e(\sigma) - \hat{e}(\sigma)) d\sigma \tag{A.3}$$

subject to

$$\hat{e} = Pf$$

$$e(\sigma), \ f(\sigma) \ g(\sigma) \quad \text{specified} \quad \sigma \in [-1,1]$$

is given by

$$P = \int_{-1}^{1} g(\sigma)e(\sigma)f^T(\sigma)d\sigma\left[\int_{-1}^{1} g(\sigma)f(\sigma)f^T(\sigma)d\sigma\right]^{-1} \qquad (A.4)$$

One way to prove (A.4) is as follows. Assume P is of maximal rank. Then we require

$$\frac{\partial\ J(P)}{\partial p^i} = \int_{-1}^{1} (e - Pf)^T g\left(-\ 2\ \frac{\partial Pf}{\partial p^i}\right)\ d\sigma = 0 \qquad (A.5)$$

where

$$P = \begin{bmatrix} p' \\ \vdots \\ p^n \end{bmatrix} \quad,\ p^i = i^{th}\ \text{row of } P\ .$$

Now

$$\frac{\partial Pf}{\partial p^i} = \begin{bmatrix} \dfrac{\partial {p^1}^T f}{\partial p^i} \\ \vdots \\ \vdots \\ \dfrac{\partial {p^n} f}{\partial p^i} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ f^T \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i^{th}\ \text{row} \qquad (A.6)$$

since

$$\frac{\partial {p^i}^T f}{\partial p^j} = f^T\ \delta_{ij}$$

Then from (A.5), (A.6),

$$\left(\int_{-1}^{1} e^T g \begin{bmatrix} f^T \\ 0 \\ \vdots \\ 0 \end{bmatrix} d\sigma \\ \vdots \\ \int_{-1}^{1} e^T g \begin{bmatrix} 0 \\ \vdots \\ 0 \\ f^T \end{bmatrix} d\sigma \right) = \left(\int_{-1}^{1} (Pf)^T \begin{bmatrix} f^T \\ 0 \\ \vdots \\ 0 \end{bmatrix} g \, d\sigma \\ \vdots \\ \int_{-1}^{1} (Pf)^T \begin{bmatrix} 0 \\ \vdots \\ 0 \\ f^T \end{bmatrix} g \, d\sigma \right)$$

or in compact form

$$\int_{-1}^{1} e \, f^T g \, d\sigma = \int_{-1}^{1} P \, f f^T g \, d\sigma$$

For a constant P

$$P = \int_{-1}^{+1} e f^T g \, d\sigma \left[\int_{-1}^{1} f \, f^T g \, d\sigma \right]^{-1}$$

|.

Theorem A-2

If $f_i(\sigma) = P_i^{(\alpha,\beta)}(\sigma)$ are Jacobi polynomials then

$$g(\sigma) = (1 - \sigma)^\alpha (1 + \sigma)^\beta$$

and

$$\int_{-1}^{1} f \, f^T g \, d\sigma = \Lambda \tag{A.7}$$

where

$$\Lambda_{ii} = \frac{2^{\alpha+\beta+1}}{(2i + \alpha + \beta + 1)i!} \cdot \frac{\Gamma(i+\alpha+1) \; \Gamma(i+\beta+1)}{\Gamma(i+\alpha+\beta+1)} \qquad (i=0,1,\ldots,d-1)$$

$$\Lambda_{ij} = 0 \; , \qquad i \neq j$$

$\Gamma(\cdot)$ is the gamma function.

Proof:

The Jacobi polynomials $f_i(\sigma) = P_i^{(\alpha,\beta)}(\sigma)$ are orthogonal on the interval $\sigma \in [-1,1]$ with respect to the weight $g(\sigma) = (1-\sigma)^\alpha (1+\sigma)^\beta$. This implies

$$\int_{-1}^1 f_i(\sigma) \, f_i(\sigma) \, g(\sigma) d\sigma = 0 \qquad i \neq j$$

and from Reference 4, p. 774

$$\int_{-1}^1 f_i^2(\sigma) g(\sigma) d\sigma = \frac{2^{\alpha+\beta+1}}{(2i+\alpha+\beta+1)i!} \cdot \frac{\Gamma(i+\alpha+1) \, \Gamma(i+\beta+1)}{\Gamma(i+\alpha+\beta+1)} \quad \Big|_{\bullet}$$

The Jacobi polynomials may be expressed by the Rodriques formula[63] p. 58,

$$P_i^{(\alpha,\beta)}(\sigma) = \frac{(-1)^i}{2^i \, i!} (1-\sigma)^{-\alpha} (1+\sigma)^{-\beta} \frac{d^i}{d\sigma^i} \left[(1-\sigma)^{\alpha+i} (1+\sigma)^{\beta+i} \right]$$

Theorem A-3

The matrix P which minimizes (A.3) subject to

$$\hat{e} = Pf$$

where $f_i(\sigma) = P_i^{(\alpha,\beta)}(\sigma)$ are Jacobi polynomials, is given approximately by

$$P_{ij} \triangleq \int_{-1}^1 g(\sigma) \, e_i(\sigma) \, f_j(\sigma) d\sigma \triangleq \sum_{k=1}^n a_k \, e_i(\sigma_k) \, f_j(\sigma_k) \qquad (A.8)$$

where σ_k is defined by

$$P_n^{(\alpha,\beta)}(\sigma_k) = 0$$

and the coefficients

$$a_k = \frac{2^{\alpha+\beta+1} \ \Gamma(n+\alpha+1) \ \Gamma(n+\beta+1)}{n! \ \Gamma(n+\alpha+\beta+1)(1-\sigma_k^2)\left[\frac{d}{d\sigma} P_n^{(\alpha,\beta)}(\sigma_k)\right]^2} \qquad (A.9)$$

The positive numbers a_k are called the Christoffel numbers. For a proof of the Gauss-Jacobi quadrature formula, (A.8), see Reference 64, p. 112.

Now that P and f are defined, the "open loop" model of the error system, as discussed in section 4.2.1, is determined by

$$\hat{e} = Pf \qquad (A.10)$$

However, to generate $\hat{e}$ from the "closed loop" model of the error system (see section 4.1.2) we must define a set of differential equations the orthogonal functions $f_i(\sigma)$ are known to obey rather than accept the "open loop" description of $\hat{e}$ as given by (A.10).

The Jacobi polynomials are eigenfunctions of the Sturm-Liouville equation (taken from p. 311 of Reference 64, Equation (7.8.4)),

$$(1-\sigma^2) \ p_i''(\sigma) \ [\beta-\alpha-(\alpha+\beta+2)\sigma] \ p_i'(\sigma) + i(i+\alpha+\beta+1) \ p_i(\sigma) = 0 \quad (A.11)$$
$$(i = 0,1,\ldots,d-1)$$

However, (A.11) is a time varying system and if placed in the first order from

$$\hat{e} = [P,0]\gamma(\sigma) \qquad (A.12)$$
$$\gamma' = D(\sigma)\gamma(\sigma) \qquad\qquad \gamma \triangleq \begin{pmatrix} p(\sigma) \\ p'(\sigma) \end{pmatrix}$$

then D is a 2d × 2d time varying matrix which complicates our intended procedure for real time state estimation of γ.

Chebyshev Error Systems

As a special case of the Jacobi polynomials, $P_i^{(\alpha,\beta)}(\sigma)$, those obtained when $\alpha = \beta = -1/2$ are called Chebyshev polynomials (multiplied by a constant, c_i).

Theorem A.4

The Chebyshev polynomials, $f_i(\sigma) = T_i(\sigma) = \frac{1}{c_i} P_i(-1/2, -1/2)(\sigma)$ yield for the $g(\sigma)$ in (A.3), (A.4)

$$g(\sigma) = (1 - \sigma^2)^{-1/2} \tag{A.13}$$

and

$$\int_{-1}^{1} f\, f^T\, g\, d\sigma = \Lambda$$

$$\Lambda = \frac{\pi}{2} \begin{bmatrix} 2 & 0 \\ 0 & I_{d-1} \end{bmatrix}, \quad \Lambda^{-1} = \frac{2}{\pi} \begin{bmatrix} 1/2 & 0 \\ 0 & I_{d-1} \end{bmatrix} \tag{A.14}$$

where

$$C_i \triangleq \frac{\Gamma(i + 1/2)}{i!\,\sqrt{\pi}} \tag{A.15}$$

The proof follows from substitution of $\alpha = \beta = -1/2$ in Theorem A.2 and from the normalization defined for the Chebyshev polynomials, resulting in the C_i in (A.15), from (22.5.31) of Reference 65.

Theorem A.5 (from Chebyshev quadrature formula,[65], Equation (25.4.38))

The matrix P which minimizes (A.3) subject to

$$\hat{e} = Pf$$

where

$$f_i(\sigma) = T_i(\sigma)$$

are Chebyshev polynomials, is given approximately by

$$\Lambda_{jj} P_{ij} = \int_{-1}^{1} g(\sigma) e_i(\sigma) f_j(\sigma) d\sigma \simeq \frac{\pi}{n} \sum_{k=1}^{n} e_i(\sigma_k) f_j(\sigma_k) \qquad (A.16)$$

where σ_k are points when

$$\sigma_k = \cos\left(\frac{2k-1}{2n} \pi\right) \qquad k=1,2,\ldots,n \qquad (A.17)$$

Equation (A.16) is the Chebyshev quadruture integration formula, (25.4.38) in Reference 65, but this result can also be obtained directly from (A.8) and (A.9), using $\alpha = \beta = -1/2$, noting (A.15) and $c_i T_i(\sigma) = P_i^{(-1/2,-1/2)}(\sigma)$. The matrix elements Λ_{jj} in (A.16) are defined by (A.14).

The Sturm-Liouville equations whose eigenfunctions are Chebyshev polynomials are given by (A.11) with $\alpha = \beta = -1/2$

$$(1-\sigma^2) p_i''(\sigma) - \sigma \, p_i'(\sigma) + i^2 \, p_i(\sigma) = 0, \qquad i=0,1,\ldots,d \qquad (A.18)$$

and the first order forms corresponding to (A.12) are

$$p_i'(\sigma) = \frac{-i\sigma}{1-\sigma^2} \, p_i(\sigma) + \frac{i}{1-\sigma^2} \, p_{i-1}(\sigma) \qquad i=0,1,2,\ldots,d \qquad (A.19)$$

The model error system using (A.19) is written

$$e = P\gamma$$
$$\hat{\gamma} = D\gamma \qquad \gamma' = \frac{d\gamma}{d\sigma} \qquad (A.20)$$

where P is given by (A.16) and

$$D = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & -\sigma & 0 & 0 & 0 \\ 0 & 2 & -2\sigma & 0 & 0 \\ 0 & 0 & 3 & -3\sigma & 0 \\ 0 & 0 & 0 & 4 & -4\sigma \\ & & & & \ddots \end{bmatrix}_{d \times d} \frac{1}{1-\sigma^2} \qquad (A.21)$$

173

Now (A.21) has a time invariant representation which can be obtained by

Theorem A.6

A time invariant model error system which can generate the Chebyshev polynomial least squares fit to the data $e(\sigma)$ is given by

$$\hat{e} = P\gamma$$

$$\gamma' = D\gamma \qquad \text{(A.22a)}$$

where P and σ_k are given by (A.16) and (A.17), respectively and

$$D = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 6 & 0 & 0 & 0 & 0 \\ 0 & 8 & 0 & 8 & 0 & 0 & 0 \\ 5 & 0 & 10 & 0 & 10 & 0 & 0 \\ 0 & 12 & 0 & 12 & 0 & 12 & 0 \\ & & & & & & & \ddots \end{bmatrix}_{d \times d} \qquad \text{(A.22b)}$$

Proof:

The only thing new to prove is that the $d \times d$ matrix D can be represented in the time invariant form of (A.22b) instead of the time varying form of (A.21). Observe that the Chebyshev polynomials can be written in the form

$$T_i(\sigma) = \sum_{k=0}^{i} C_k \, \sigma^k \qquad \text{(A.23)}$$

and the inverse relationships

$$\sigma^i = b_1^{-1} \sum_{k=0}^{i} d_k T_k(\sigma) \qquad\qquad (A.24)$$

can be taken from Table A-1. Differentiating (A.23) with respect to σ and then making the substitution (A.24) into (A.23) gives in a straightforward manner

$$T' = DT \qquad\qquad T^T \triangleq (T_o, T_1, \ldots, T_{d-1}) \ , \ T' \triangleq \frac{dT}{d\sigma}$$

where D is given by (A.22b)

Note that when normalized time, σ, is replaced by t we have from (A.2) $\dot{\gamma} = (2/\tau)\gamma'$ and the matrix corresponding to $\dot{\gamma} = D\gamma$ is obtained by multiplying (A.22b) by $2/\tau$.

Now we proceed to see how any solution of (A.22a) relates to the Chebyshev polynomials, $f_i(\sigma) = T_i(\sigma)$.

<u>Theorem A.7</u>

Any solution of (A.22a) satisfies

$$\hat{e} = P(\gamma(o))T \qquad\qquad T^T \triangleq \Big(T_o(\sigma), \ T_1(\sigma), \ldots, T_{d-1}(\sigma)\Big) \qquad (A.25)$$

where the $T_i(\sigma)$ are Chebyshev polynomials of degree i and the coefficients of $T(\sigma)$ are functions of the initial conditions of (A.22a), specifically,

$$P \ \gamma(0) \ \triangleq P \begin{bmatrix} \gamma^T(0) & Q_4 \\ \gamma^T(0) & Q_2 \\ & \vdots & \\ \gamma^T(0) & Q_d \end{bmatrix} \qquad\qquad (A.26)$$

where

Table A-1 (from Reference 5, p. 795)

COEFFICIENTS OF THE CHEBYSHEV POLYNOMIALS AND THEIR INVERSE

	σ^0	σ^1	σ^2	σ^3	σ^4	σ^5	σ^7
b_i	1	1	2	4	8	16	32
T_0	1 1		1		3		10
T_1		1 1		3		10	
T_2	-1		2 1		4		15
T_3		-3		4 1		5	
T_4	-1		-8		8 1		6
T_5		5		-20		16 1	
T_6	-1		-18		-48		32 1

Example: $T_6(\sigma) = 32\,\sigma^6 - 48\,\sigma^4 + 18\,\sigma^2 - 1$

$$\sigma^6 = \frac{1}{32}\left[10\,T_0 + 15\,T_2 + 6\,T_4 + T_6\right]$$

$$Q_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ & & & & \ddots \end{bmatrix}_{d \times d}$$

$$Q_2 = \begin{bmatrix} 1 & 1 & \\ 1 & 0 & 0 \\ \hline / & 0 & 0 \end{bmatrix}_{d \times d}$$

$$Q_3 = \begin{bmatrix} 341 & \\ 440 & 0 \\ 100 & \\ \hline 0 & 0 \end{bmatrix}_{d \times d}$$

$$Q_4 = \begin{bmatrix} 13 & 18 & 6 & 1 & \\ 18 & 24 & 6 & 0 & \\ 6 & 6 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & \\ \hline & & 0 & & 0 \end{bmatrix}_{d \times d}$$

$$\vdots$$

Proof:

The state transition matrix for A in (A.22b) may be computed exactly from the finite series

$$\Phi(\sigma, \sigma_o) = I_d + D\hat{\sigma} + \frac{1}{2} D^2 \hat{\sigma}^2 + D^3 \frac{\hat{\sigma}^3}{3!} + \text{---} D^{d-1} \frac{\hat{\sigma}^{d-1}}{(d-1)!} \qquad (A.27)$$

$$\hat{\sigma} \triangleq \sigma - \sigma_o$$

Then, from (A.22b), (A.27) becomes

$$\Phi(\sigma,\sigma) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \hat{\sigma} & 1 & 0 & 0 & 0 \\ 2\hat{\sigma}^2 & 4\hat{\sigma} & 1 & 0 & 0 \\ 3\hat{\sigma} + 4\hat{\sigma}^3 & 12\hat{\sigma}^2 & 6\hat{\sigma} & 1 & 0 \\ 16\hat{\sigma}^2 + 8\hat{\sigma}^4 & 8\hat{\sigma} + 32\hat{\sigma}^3 & 24\hat{\sigma}^2 & 8\hat{\sigma} & 1 \\ & & & & \ddots \end{bmatrix}$$

By substituting $\hat{\sigma} = \sigma + 1$ (taking $\sigma_o = -1$) and using (A.24) and Table A-1, $\Phi(\sigma, -1)$ can be written, after some labor,

$$\Phi(\sigma,-1)=\begin{bmatrix} T_o & 0 & 0 & 0 \\ T_o+T_1 & T_o & 0 & 0 \\ 3T_o+4T_1+T_2 & 4T_o+4T_1 & T_o & 0 \\ 13T_o+18T_1+6T_2+T_3 & 18T_o+24T_1+6T_2 & 6T_o+6T_1 & T_o \\ 59T_o+88T_1+36T_2+8T_3+T_4 & 88T_o+128T_1+48T_2+8T_3 & 36T_o+48T_1+12T_2 & 8T_o+8T_1 \\ & & & \ddots \end{bmatrix}$$

or,

$$\Phi(\sigma,-1) = \begin{bmatrix} \phi_1^T \\ \phi_2^T \\ \vdots \\ \phi_d^T \end{bmatrix} \tag{A.28}$$

where

178

$$\phi_1 = Q_1 T \triangleq \begin{bmatrix} 1 & 0 \\ \hline 0 & 0 \end{bmatrix} \begin{pmatrix} T_o \\ \vdots \\ T_{d-1} \end{pmatrix}$$

$$\phi_2 = Q_2 T \triangleq \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & \\ \hline & 0 & 0 \end{bmatrix} \begin{pmatrix} T_o \\ \vdots \\ T_{d-1} \end{pmatrix}$$

$$\phi_3 = Q_3 T \triangleq \begin{bmatrix} 3 & 4 & 1 & \\ 4 & 4 & 0 & 0 \\ 1 & 0 & 0 & \\ \hline & 0 & & 0 \end{bmatrix} \begin{pmatrix} T_o \\ \vdots \\ T_{d-1} \end{pmatrix}$$

$$\phi_4 = Q_4 T = \begin{bmatrix} 13 & 18 & 6 & 1 & \\ 18 & 24 & 6 & 0 & \\ 6 & 6 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & \\ \hline & 0 & & & 0 \end{bmatrix} \begin{pmatrix} T_o \\ \vdots \\ T_{d-1} \end{pmatrix}$$

$$\phi_5 = Q_5 T = \begin{bmatrix} 59 & 88 & 36 & 8 & 1 & \\ 88 & 128 & 48 & 8 & 0 & \\ 36 & 48 & 12 & 0 & 0 & 0 \\ 8 & 8 & 0 & 0 & 0 & \\ 1 & 0 & 0 & 0 & 0 & \\ \hline & & 0 & & & 0 \end{bmatrix} \begin{pmatrix} T_o \\ \vdots \\ T_{d-1} \end{pmatrix}$$

Now since

$$\gamma(\sigma) = \Phi(\sigma,-1)\,\gamma(0) = \begin{pmatrix} \phi_1^T \gamma(0) \\ \vdots \\ \phi_d^T \gamma(0) \end{pmatrix} = \begin{pmatrix} \gamma^T(0)\,\phi_1 \\ \vdots \\ \gamma^T(0)\,\phi_d \end{pmatrix} \qquad (A.29)$$

it follows from $\hat{e} = P\gamma$, (A.29) and (A.28) that

$$\hat{e}(\sigma) = P(\gamma(0))T$$

where

179

$$P = P \begin{bmatrix} \gamma^T(0) & Q_1 \\ & \vdots \\ \gamma^T(0) & Q_d \end{bmatrix}$$

On the basis of Theorem A.7 the closed loop "synthetic mode" description of the error system (discussed in section ___ and described by (A.22a), is seen to have certain advantages over open loop descriptions (a priori selection of $\hat{e}(\sigma)$ as a prescribed function of σ). Specifically when on-line state estimation of $\gamma(0)$ is accomplished (in response to real measurements) then (A.25) shows that the model error vector, $\hat{e}(\sigma)$, is a least squares fit to the actual error vector $e(\sigma)$ using Chebyshev polynomials whose coefficients are appropriately and automatically selected by the current state estimates for $\gamma(\sigma)$.

Fourier Error Systems

Let the Chebyshev polynomials be written

$$T_n(\sigma) = \cos(n \cos^{-1} \sigma) \qquad \sigma \in [-1,1] \qquad \text{(A.30)}$$

and make the change of variable

$$\cos \theta = \sigma \qquad \theta \in [\pi.0]$$

then,

$$F_n(\theta) = T_n(\sigma) = \cos n\theta \qquad \text{(A.31)}$$

Furthermore make the change of variable

$$\theta = wt + \psi \qquad t \in [0,\tau] \qquad \text{(A.32)}$$

For consistant limits

$$\cos \psi = -1 \quad \Rightarrow \quad \psi = k\pi \quad , \quad k = \pm 1,3,5 \text{ ----}$$

$$\cos(\omega\tau + \psi) = 1 \quad \Rightarrow \quad \tau = \frac{k\pi}{w}$$

Then

$$F_n(t) = \cos(n(\omega t + \pi))$$

and Appendix A may be used substituting for σ

$$\sigma = \cos(wt + \pi) = -\cos wt \qquad \sigma \in [-1,1]$$

$$\theta = wt + \pi \qquad t \in [0, k\pi/w] \qquad \text{(A.33)}$$

where k is any odd integer. Now to compute the D matrix corresponding to the choice $F_n(wt + \pi)$ we note $\frac{d\sigma}{dt} = w \sin wt$ and

$$\dot{\gamma} = \gamma' \frac{d\sigma}{dt}$$

Then from (A.33)

$$\dot{\gamma} = \frac{d\sigma}{dt} \, D\gamma = \omega \sin \omega t \, D\gamma \qquad \text{(A.34)}$$

Thus, the correct D matrix corresponding to Fourier series in the time varying choice

$$D(t) = w \sin wt \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 3 & 0 & 6 & 0 \\ & & & & \ddots \\ & & & & & \ddots \end{bmatrix}_{dd} \qquad \text{(A.35)}$$

A time invariant characterization for the Fourier error system is possible by increasing the order of the error system to 2d.

181

APPENDIX B: EVALUATION OF THE COST WITH RESPECT TO MODEL $\mathcal{d}_1$

In this appendix we wish to derive an expression for

$$J(T_{21},d) = E \int_0^T y^{o^T}(t) \ Q \ y^o(t)dt \tag{B.1}$$

subject to

$$\dot{x} = Ax + G_r \ w_r + G_d \ w_d \tag{B.2}$$

$$y^o = y + \tilde{y} = cx + \tilde{y}$$

$$E[x_o] = 0 \qquad = \bar{x}_o$$

$$E[x_o x_o^T] = \Sigma_o$$

$$E[W_r] = 0 \qquad = \bar{w}_r$$

$$E[W_r(t)w_r(\sigma)^T] = Q_r \ \delta(t-\sigma) \qquad (\delta \text{ is Dirac delta})$$

$$E[x_o \ w_r^T(t)] = 0$$

$$E[W_d] = w_d \qquad = \bar{w}_d$$

$$E[w_d \ w^T] = w_d \ w_d^T$$

$$E[\tilde{y}] \qquad = 0 \qquad = \bar{\tilde{y}}^1$$

$$E[\tilde{y}(t) \ \tilde{y}(\sigma)^T] = R_y \ \delta(t-\sigma)$$

$$E[\tilde{y}(t) \ w_r(t)^T] = 0$$

$$E[\tilde{y}(t) \ x_o^t] \qquad = 0$$

Define

$$\Sigma \triangleq [(x-\bar{x})(x-\bar{x})^T] \ , \quad \bar{x} \triangleq E[x(t)]$$

The form of the solution for (B.2) is

$$x(t) = \Phi(t,t_o)x_o + \int_0^t \Phi(t,\tau)(G_r w_r + G_d w_d)d\sigma \tag{B.3}$$

...DING PAGE BLANK NOT FI...

Then

$$\bar{x}(t) = \int_0^t \Phi(t,\sigma) \ G_d \bar{w}_d \ d\sigma \tag{B.4}$$

Now the covariance matrix $\sum$ becomes

$$\sum = E\left\{ \left[\Phi(t,t_o)x_o + \int_0^t \Phi(t,\sigma)G_r w_r d\sigma \right] \left[\Phi(t,t_o)x_o + \int_0^t \Phi(t,\sigma)G_r w_r d\delta \right]^T \right\}$$

But since w_r is assumed white, the screening property of the Dirac delta function present in the covariance kernel for $w_r(t)$, see under (B.2), allows us to write

$$\sum = \Phi(t,0) \sum_o \Phi(t,0) + \int_0^t \Phi(t,\sigma)G_r Q_r G_r^T \ \Phi^T(t,\sigma)d\sigma \tag{B.5}$$

Differentiating (B.5) with respect to time yields

$$\dot{\sum} = A\sum + \sum A^T + G_r \ Q_r \ G_r^T \ , \qquad \sum(0) = \sum_o \tag{B.6}$$

An alternate expression for (B.1) is

$$J(T_{21},d) = E \int_0^T (cx + \tilde{y})^T \ Q(cx + \tilde{y})$$

$$= E \int_0^T (2\tilde{y}^T Qcx + \tilde{y}^T Q\tilde{y} + x^T c^T Qcx)dt \tag{B.7}$$

But from the fact that the state is uncorrelated with the measurement error $\tilde{y}$, and using the trace notation

$$J(T_{21},d) = tr \int_0^T (Q \ R_y(t) + c^T Qc \ E[x(t)x^T(t)])dt \tag{B.8}$$

Now since

$$\sum = E[(x-\bar{x})(x-\bar{x})^T] = E[xx^T] - \bar{x} \ \bar{x}^T$$

Then

$$E[x \, x^T] = \sum + \bar{x} \, \bar{x}^T$$

allowing (B.8) to be written

$$J(T_{21}, d) = \text{tr} \int_0^T [QR_y(t) + C^T QC\Sigma + C^T QC \, \bar{x} \, \bar{x}^T] dt \qquad (B.9)$$

But since R_y does not depend upon T_{21}, an equivalent problem is

$$J(T_{21}) = \text{tr} \, C^T QC \int_0^T (\Sigma + \bar{x} \, \bar{x}^T) dt \qquad (B.10)$$

APPENDIX C: FLEXIBLE SPACECRAFT WITH THERMAL DEFLECTIONS

For our present purposes LST is modeled as a central rigid body to which is attached appendages such as the solar panels, light shield, and possibly the Service Support Module (SSM). The central rigid body, B, is the primary/secondary mirror assembly together with either all or part of the SSM and its subassemblies.
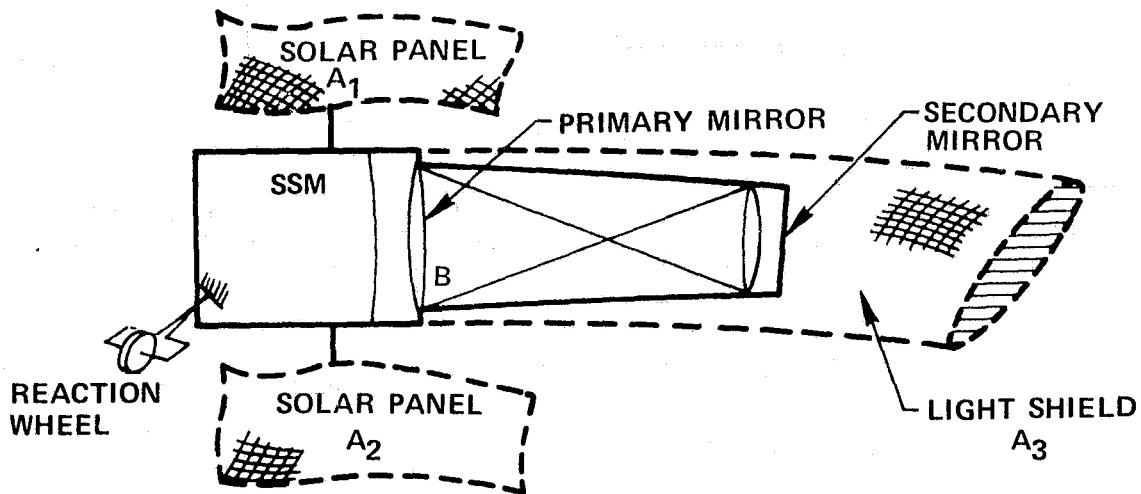
We require the rigid body B to maintain its inertial attitude in pitch and yaw within .005 seconds of arc over a five hour period.

Following[66], $\underline{R}$ is the inertial position of the system mass center, M is the system mass, and the applied resultant force is the vector $\underline{F}$, see Figure C-1.

The procedure and notation of Reference 66 will be followed without elementary developments. For further details the reader should see section III of Reference 66. The addition of thermally

PRECEDING PAGE BLANK NOT FILMED

$\bar{t}$ = TIME OF SOLAR
INCIDENCE

$\underline{\Delta}^i$ = CHANGE IN EQUILIBRIUM
POSITION OF A^i, $t > \bar{t}$

SUBBODY A^i

INERTIAL
REFERENCE, $\underline{\ell}$
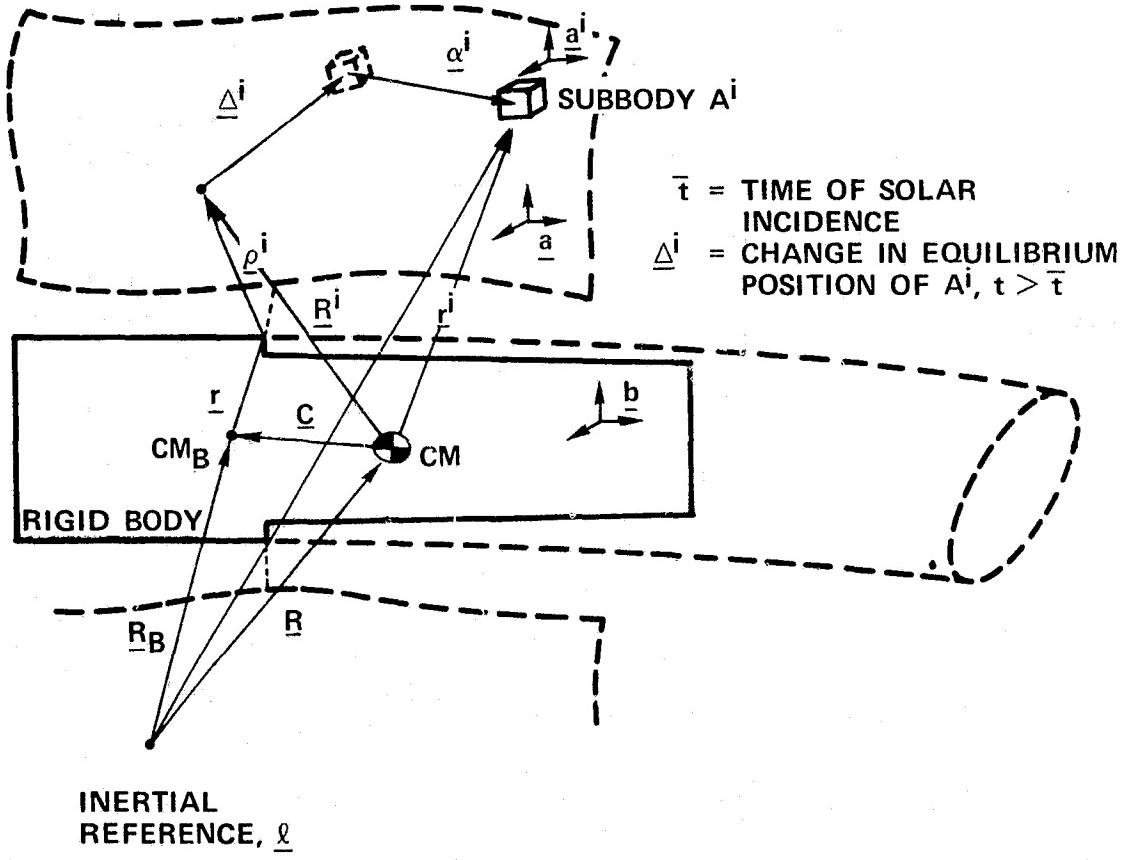
Figure C-1. Flexible Body With Thermal Deflections.

induced changes in the undeformed position of the appendage is the only fundamental difference from the development in Reference 66, section IIIA. The plan is to write 6N + 6 scalar equations (2N + 2 Gibbsian vector equations).

$$\underline{F}_B = M_B \ddot{\underline{R}}_B$$
$$\underline{T}_B = \dot{\underline{H}}_B$$

$\left.\right\}$ rigid body B

$$\underline{F}^i = m^i \ddot{\underline{r}}^i$$
$$\underline{T}^i = \dot{\underline{H}}^i$$

, i=1-N

$\left.\right\}$ appendage (N discrete nodal bodies)

The resulting system of scalar equations will be linearized for further study and truncation. Preliminary facts and definitions which will be useful are: (see Figure C-1)

$$\underline{b} = \Theta \, \underline{\ell}$$

$$\underline{a} = C \, \underline{b} \quad , \quad C = I_3 \text{ indentity matrix}$$

$$\underline{a}^i = B \, \underline{a} \quad , \quad B = I - \tilde{\beta}^i$$

$$\underline{\omega} \triangleq {}^{\ell}\underline{\omega}^b = \text{angular velocity of B with respect to inertial space.}$$

$${}^{\ell}\underline{\omega}^i = {}^{a}\underline{\omega}^i + \underline{\omega} = \dot{\underline{\beta}}^i + \underline{\omega}$$

$${}^{a}\underline{\omega}^i = \{a\}^T \dot{\underline{\beta}}^i = \dot{\underline{\beta}}^i$$

$$M_B \, \underline{c} + \sum_{i=1}^{N} (\underline{c} + \underline{\lambda} + \underline{\rho}^i + \underline{\Delta}^i + \underline{\alpha}^i) m^i = 0 \qquad \text{(CM definition)}$$

$$\ddot{\underline{R}} = 0 \Rightarrow \ddot{\underline{c}} = \ddot{\underline{R}}_B \qquad \text{(CM non-accelerating)}$$

$$\frac{{}^{\ell}d}{dt} \underline{V} = \frac{{}^{b}d}{dt} \underline{V} + {}^{\ell}\underline{\omega}^b \times \underline{V}$$

$$\frac{{}^{\ell}d^2}{dt^2} \underline{V} = \frac{{}^{b}d^2}{dt^2} \underline{V} + \left(2 {}^{\ell}\underline{\omega}^b\right) \times \left(\frac{{}^{b}d}{dt} \underline{V}\right) + \left({}^{\ell}\underline{\omega}^b\right) \times \left({}^{\ell}\underline{\omega}^b \times \underline{V}\right) + \left(\frac{{}^{\ell}d}{dt} {}^{\ell}\underline{\omega}^b\right) \times (\underline{V})$$

Appendage Equations:

Write, from Newton's second law,

$$\underline{F}^i = m^i\,(\ddot{\underline{R}}^{\,0} + \ddot{\underline{c}} + \ddot{\underline{r}} + \ddot{\underline{\Delta}}^i + \ddot{\underline{\alpha}}^i + \ddot{\underline{\rho}}^i) \tag{C.1}$$

But from CM definition and $\ddot{\underline{c}} = \ddot{\underline{R}}_B$, write,

$$\mathcal{M}\,\ddot{\underline{R}}_B + \sum_{i=1}^{N}(\ddot{\underline{r}} + \ddot{\underline{\rho}}^i + \ddot{\underline{\Delta}}^i + \ddot{\underline{\alpha}}^i)m^i = 0 \tag{C.2}$$

and (C.1) becomes, using (C.2),

$$\underline{F}^i = m^i(\ddot{\underline{r}} + \ddot{\underline{\Delta}}^i + \ddot{\underline{\alpha}}^i + \ddot{\underline{\rho}}^i) - \frac{m^i}{\mathcal{M}}\sum_{k=1}^{N}\left[(\ddot{\underline{r}} + \ddot{\underline{\rho}}^k + \ddot{\underline{\Delta}}^k + \ddot{\underline{\alpha}}^k)m^k\right] \tag{C.3}$$

or

$$\underline{F}^i = \left(1 - \frac{\mathcal{M}^A}{\mathcal{M}}\right)\ddot{\underline{r}}\,m^i + \left(1 - \frac{m^i}{\mathcal{M}}\right)m^i(\ddot{\underline{\Delta}}^i + \ddot{\underline{\rho}}^i + \ddot{\underline{\alpha}}^i)$$

$$-\frac{m^i}{\mathcal{M}}\left[\sum_{\substack{k=1\\k\neq i}}^{N}(\ddot{\underline{\rho}}^k + \ddot{\underline{\Delta}}^k + \ddot{\underline{\alpha}}^k)m^k\right] \tag{C.4}$$

$\mathcal{M}^A \triangleq$ mass of appendage.

Now compute

$$\ddot{\underline{r}} = \overset{\circ\circ}{\underline{r}} + 2\,\underline{\omega}\times\dot{\underline{r}} + \underline{\omega}\times(\underline{\omega}\times\underline{r}) + \dot{\underline{\omega}}\times\underline{r}$$

$$\ddot{\underline{r}} = \dot{\underline{\omega}}\times\underline{r}\quad\text{(linearized)} \tag{C.5}$$

Now

$$\ddot{\underline{\rho}}^i = \ddot{\underline{\rho}}^{i\,0} + 2\,^\ell\underline{\omega}^a\times\dot{\underline{\rho}}^{i\,0} + {}^\ell\underline{\omega}^a\times\left({}^\ell\underline{\omega}^a\times\underline{\rho}^i\right) + {}^\ell\dot{\underline{\omega}}^a\times\underline{\rho}^i$$

$$\ddot{\underline{\rho}}^i = \dot{\underline{\omega}}\times\underline{\rho}^i\qquad\text{(linearized)} \tag{C.6}$$

Now

$$\underline{\ddot{\Delta}}^i + \underline{\ddot{\alpha}}^i = \overset{\circ\circ}{\underline{\Delta}}{}^i + \overset{\circ\circ}{\underline{\alpha}}{}^i + 2\,{}^\ell\underline{\omega}{}^{a^i} \times (\overset{\circ}{\underline{\Delta}}{}^i + \overset{\circ}{\underline{\alpha}}{}^i) + {}^\ell\underline{\omega}{}^{a^i} \times ({}^\ell\underline{\omega}{}^{a^i} \times (\overset{\circ}{\underline{\Delta}}{}^i + \underline{\alpha}{}^i))$$

$$+ {}^\ell\underline{\dot{\omega}}{}^{a^i} \times (\underline{\Delta}{}^i + \underline{\alpha}{}^i)$$

$$= \overset{\circ\circ}{\underline{\Delta}}{}^i + \overset{\circ\circ}{\underline{\alpha}}{}^i \qquad \text{(linearized)} \qquad (C.7)$$

Then (C.4) becomes

$$\underline{F}^i = \left(1 - \frac{\mathcal{M}^A}{\mathcal{M}}\right) m^i\,\underline{\dot{\omega}} \times \underline{r} + \left(1 - \frac{m^i}{\mathcal{M}}\right) m^i (\underline{\dot{\omega}} \times \underline{\rho}^i + \overset{\circ\circ}{\underline{\Delta}}{}^i + \overset{\circ\circ}{\underline{\alpha}}{}^i)$$

$$- \frac{m^i}{\mathcal{M}} \left[\sum_{\substack{k=1 \\ k \neq i}}^{N} m^k (\underline{\dot{\omega}} \times \underline{\rho}^k + \underline{\ddot{\Delta}}{}^k + \underline{\ddot{\alpha}}{}^i)\right] \qquad (C.8)$$

The Euler equations for the appendage sub-body A^i are

$$\underline{T}^i = \underline{\dot{H}}{}^i$$

where

$$\underline{\dot{H}}{}^i = \underline{\underline{\mathbb{I}}}{}^i \cdot {}^\ell\underline{\omega}{}^i + \underline{\underline{\mathbb{I}}}{}^i \cdot {}^\ell\underline{\dot{\omega}}{}^i$$

$$= \left(\overset{\circ}{\underline{\underline{\mathbb{I}}}}{}^i + {}^\ell\underline{\omega}{}^i \times \underline{\underline{\mathbb{I}}}{}^i\right) \cdot {}^\ell\underline{\omega}{}^i + \underline{\underline{\mathbb{I}}}{}^i \cdot {}^\ell\underline{\dot{\omega}}{}^i$$

$$= {}^\ell\underline{\omega}{}^i \times \underline{\underline{\mathbb{I}}}{}^i \cdot {}^\ell\underline{\omega}{}^i + \underline{\underline{\mathbb{I}}}{}^i \cdot {}^\ell\underline{\dot{\omega}}{}^i$$

$$\underline{\dot{H}}{}^i = \underline{\mathbb{J}}{}^i \cdot ({}^\ell\underline{\dot{\omega}}{}^b + {}^a\underline{\dot{\omega}}{}^i) \qquad \text{(Linearized)} \qquad (C.9)$$

But

$${}^a\underline{\dot{\omega}}{}^i = {}^a\overset{\circ}{\underline{\omega}}{}^i + {}^a\underline{\omega}{}^i \times {}^a\underline{\omega}{}^i = {}^a\overset{\circ}{\underline{\omega}}{}^i = \overset{\circ\circ}{\underline{\beta}}{}^i$$

Hence

$$\underline{T}^i \approx \underline{\underline{\mathbb{J}}}{}^i \cdot (\underline{\dot{\omega}} + \overset{\circ\circ}{\underline{\beta}}{}^i) \qquad (C.10)$$

Now for matrix forms we record

$$\underline{\omega} = \{b\}^T \omega$$

$$\underline{\dot{\omega}} = \{b\}^T \dot{\omega}$$

$$\ddot{\underline{\beta}}^i = \{a^i\}^T \ddot{\beta}^i = \{b\}^T \ddot{\beta}^i \qquad \text{(for small } \beta\text{)}$$

$$\underline{r} = \{b\}^T r$$

$$\underline{\rho}^i = \{a\}^T \rho^i = \{b\}^T \rho^i \qquad \text{(for small } \beta\text{)}$$

$$\ddot{\underline{\Delta}}^k = \{a^i\}^T \ddot{\Delta}^K = \{b\}^T \ddot{\Delta}^k \qquad \text{(for small } \Delta^k, \alpha^k, \beta^k\text{)}$$

$$\ddot{\underline{\alpha}}^k = \{a^i\}^T \ddot{\alpha}^k = \{b\}^T \ddot{\alpha}^k \qquad \text{(for small } \Delta^k, \alpha^k, \beta^k\text{)}$$

Now (C.8) becomes, in frame $\underline{b}$, (note $\underline{F}^i = \{b\}^T F^i$)

$$F^i = \left(1 - \frac{\mathcal{M}^A}{\mathcal{M}}\right) m^i \, \tilde{\omega} r + \left(1 - \frac{m^i}{\mathcal{M}}\right) m^i (\tilde{\omega} \, \rho^i + \ddot{\Delta}^i + \ddot{\alpha}^i)$$

$$- \frac{m^i}{\mathcal{M}} \left[\sum_{\substack{k=1 \\ k \neq i}}^{N} (\tilde{\omega} \, \rho^k + \ddot{\Delta}^k + \ddot{\alpha}^k) m^k \right] \qquad\qquad \text{(C.11)}$$

and (C.10) becomes (note $\underline{\underline{I}}^i = \{a^i\}^T I^i \{a^i\} \approx \{b\}^T I^i \{b\}$ for small $\Delta^k, \alpha^k, \beta^k$)

$$T^i = J^i (\dot{\omega} + \ddot{\beta}^i) \qquad\qquad \text{(C.12)}$$

Now combine (C.11), (C.12) using the definitions

$$q^T \triangleq (\alpha^T, \, \beta^T) \qquad\qquad \alpha^T \triangleq \left(\alpha^{1^T}, \, \alpha^{2^T}, \ldots, \alpha^{N^T} \right)$$

$$\alpha^{i^T} = \left(\alpha_1^i, \, \alpha_2^i, \, \alpha_3^i \right), \qquad i = 1, N$$

$$\beta^T = \left(\beta^{1^T}, \, \beta^{2^T}, \ldots, \beta^{N^T} \right)$$

$$\beta^{i^T} = \left(\beta_1^i, \, \beta_2^i, \, \beta_3^i \right)$$

$$T^i = \overline{T}^i + \hat{T}^i$$

$$F^i = \overline{F}^i + \hat{F}^i$$

$$\overline{T}^i = - \, {}^t K^i \, \beta^i \quad \text{(interconnection moments between subbodies } A^i\text{)}$$

$$\overline{F}^i = - \, {}^f K^i \, \alpha^i \quad \text{(interconnection forces between subbodies } A^i\text{)}$$

192

t_K^i = torsional stiffness (3×3), f_K^i = translational stiffness (3×3) and where $\hat{F}^i$, $\hat{T}^i$, are externally applied forces and torques on sub-body A^i. Rewriting (C.11)

$$-\frac{m^i}{\mathcal{M}} m^i \ddot{\alpha}^1 \dots + \left(1-\frac{m^i}{\mathcal{M}}\right) m^i \ddot{\alpha}^i \dots - \frac{m^i}{\mathcal{M}} m^N \ddot{\alpha}^N + {}^f K^i \alpha^i$$

$$= \hat{F}^i + \left[m^i (\tilde{r} + \tilde{\rho}^i) - \frac{m^i}{\mathcal{M}} \sum_{k=1}^{N} m^k (\tilde{r} + \tilde{\rho}^k) \right] \dot{\omega} +$$

$$m^i \ddot{\Delta}^i - \frac{m^i}{\mathcal{M}} \sum_{k=1}^{N} m^k \ddot{\Delta}^k \qquad (C.13)$$

For N equations (i=1 --- N) the matrix form can be written

$$^f_M \ddot{\alpha} + {}^f_K \alpha = \hat{f} + {}^f_M \tilde{R} \dot{\omega} + {}^f_M \ddot{\Delta} \qquad (C.14)$$

where

$$^f_{\underline{\underline{M\Delta}}} \begin{bmatrix} m^1\left(1-\frac{m^1}{\mathcal{M}}\right) I_3 & -\frac{m^1}{\mathcal{M}} m^2 I_3 - - - - & -\frac{m^i}{\mathcal{M}} m^N I_3 \\ -\frac{m^2 m^1}{\mathcal{M}} I_3 & m^2\left(1-\frac{m^2}{\mathcal{M}}\right) I_3 - - - - & -\frac{m^2 m^N}{\mathcal{M}} I_3 \\ \vdots & & \\ & & -\frac{m^{N-1} m^N}{\mathcal{M}} I_3 \\ -\frac{m^N m^1}{\mathcal{M}} I_3 & -\frac{m^N m^2}{\mathcal{M}} I_3 - - - - & m^N\left(1-\frac{m^N}{\mathcal{M}}\right) I_3 \end{bmatrix}_{3N \times 3N}, \quad I_3 \underline{\underline{\triangle}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$^f_{\underline{\underline{K\Delta}}} \begin{bmatrix} {}^f K^1 & & 0 \\ & \ddots & \\ 0 & & {}^f K^N \end{bmatrix}_{3N \times 3N}$$

193

$$\ddot{\Delta} \triangleq \begin{pmatrix} \ddot{\Delta}^1 \\ | \\ | \\ | \\ \ddot{\Delta}^N \end{pmatrix}_{3N \times 1} \qquad , \qquad \tilde{r} \triangleq \begin{bmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{bmatrix}_{3 \times 3}$$

$$f = \begin{pmatrix} \hat{F}^1 \\ | \\ | \\ | \\ \hat{F}^N \end{pmatrix}_{3N \times 1}$$

$$\tilde{R} \triangleq \begin{bmatrix} \tilde{R}^1 \\ \tilde{R}^2 \\ | \\ | \\ \tilde{R}^N \end{bmatrix} \qquad , \qquad R^i = r + \rho^i$$

Now, from (C.12),

$$J \ddot{\beta} + {}^t K \beta = \hat{T} - J' \dot{\omega} \tag{C.15}$$

where

$$J = \begin{bmatrix} J^1 & & 0 \\ & \diagdown & \\ 0 & & J^N \end{bmatrix}_{3N \times 3N} \qquad J' = \begin{bmatrix} J^1 \\ | \\ J^N \end{bmatrix}_{3N \times 3} , \qquad \begin{array}{l} J^i = \text{inertia matrix of} \\ \text{subbody } A^i \text{ about} \\ \text{its mass center} \end{array}$$

$${}^t K = \begin{bmatrix} {}^t K^1 & & 0 \\ & \diagdown & \\ 0 & & {}^t K^N \end{bmatrix}_{3N \times 3N}$$

Together (C.14) and (C.15) imply

$$M \ddot{q} + K q = L' \tag{C.16}$$

where

194

$$M = \begin{bmatrix} {}^f M & 0 \\ 0 & J \end{bmatrix} \quad , \quad K = \begin{bmatrix} {}^f K & 0 \\ 0 & {}^t K \end{bmatrix}$$

$$L' = \ell - \Gamma^T \dot{\omega} + {}^f\mathcal{M} \ddot{\Delta} \qquad (C.17)$$

$$\ell \triangleq \begin{pmatrix} \hat{f} \\ \hat{T} \end{pmatrix}_{6N \times 1} \quad , \quad \Gamma^T \triangleq \begin{bmatrix} {}^f-M\,R \\ +J' \end{bmatrix}_{6N \times 3} \quad , \quad {}^f\mathcal{M} \triangleq \begin{bmatrix} {}^f M \\ 0 \end{bmatrix}_{6N \times 3N}$$

For appendage equation summary, see Figure C-2.

Now the thermally induced displacements of the appendage subbodies are assumed to satisfy the first order descriptions

$$\Delta = \left[I - e^{-\Lambda(t-\bar{t})} \right] \Delta_o \qquad (C.18)$$

where

$$e^{-\Lambda(t-\bar{t})} = \begin{bmatrix} e^{-\Lambda^1(t-\bar{t})} & & 0 \\ & \diagdown & \\ 0 & & e^{-\Lambda^N(t-\bar{t})} \end{bmatrix}_{3N \times 3N}$$

$$e^{-\Lambda^i(t-\bar{t})} \triangleq \begin{bmatrix} e^{-\lambda^i_1(t-\bar{t})} & 0 & 0 \\ 0 & e^{-\lambda^i_2(t-\bar{t})} & 0 \\ 0 & 0 & e^{-\lambda^i_3(t-\bar{t})} \end{bmatrix}_{3 \times 3}$$

$\left(\lambda^i_k \right)^{-1}$ = thermal time constant associated with the kth direction for subbody A^i (assumed given)

Δ_o = steady state displacement due to change in temperature

195

For elements of the appendage, A^i, which are assumed unaffected by solar incidence, set $\lambda^j \equiv 0$. If all such thermal effects are to be ignored set $\lambda^j \equiv 0$, $j=1-N$, and note that $\Delta \equiv 0 \equiv \ddot{\Delta}$ and (16), (17) reduce to

$$M\ddot{q} + K\,q = \ell - \Gamma^T \dot{\omega} \qquad (C.19)$$

Now to obtain the expression for $\ddot{\Delta}$ needed in (C.17), differentiate (C.18) twice to get

$$\ddot{\Delta} = -\Lambda^2\, e^{-\Lambda(t-\bar{t})} \Delta_o \qquad (C.20)$$

where

$$\Lambda^2 = \Lambda\,\Lambda \quad , \quad \Lambda = \begin{bmatrix} \Lambda^1 & & 0 \\ & \diagdown & \\ 0 & & \Lambda^N \end{bmatrix} , \quad \Lambda^i = \begin{bmatrix} \lambda^i_1 & 0 & 0 \\ 0 & \lambda^i_2 & 0 \\ 0 & 0 & \lambda^i_3 \end{bmatrix}$$

If we further assume that the thermal time constant associated with each modal body and each direction is the same then

$$\ddot{\Delta} = -\lambda^2\, e^{-\lambda(t-\bar{t})} \Delta_o \qquad (C.21)$$

and the appendage Equations (C.16) become

$$M\ddot{q} + Kq = \ell - \Gamma^T \dot{\omega} - f_M\, \lambda^2 e^{-\lambda(t-\bar{t})} \Delta_o \qquad (C.22)$$

If the system is time invariant then the convenient choice $\bar{t} = 0$ is allowed. The appendage equations are summarized in Figure C-2.
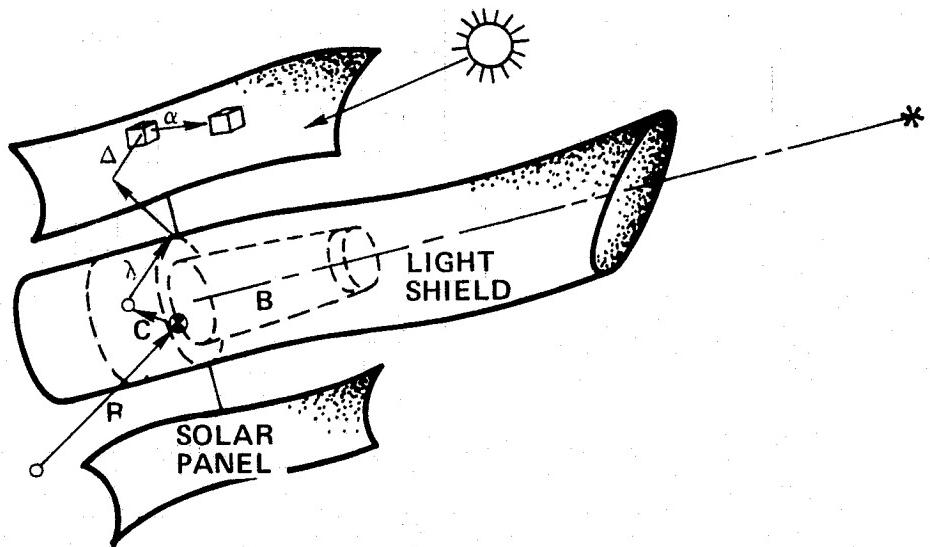
$$M\ddot{q} + G\,\dot{q} + K\,q = \ell - \Gamma^T\,\dot{\omega} + {}^f_M\ddot{\Delta}$$

$$
{}^f\!\mathcal{M} = \begin{bmatrix} {}^f_M \\ 0 \end{bmatrix}, \quad
M = \begin{bmatrix} {}^f_M & 0 \\ 0 & J \end{bmatrix}_{6N\times6N}, \quad
{}^f_M = \begin{bmatrix}
m'\left(1-\dfrac{m^1}{\mathcal{M}}\right)I_3 & -\dfrac{m^1 m^2}{\mathcal{M}}I_3 & \cdots & -\dfrac{m^1 m^N}{\mathcal{M}}I_3 \\
-\dfrac{m^2 m^1}{\mathcal{M}}I_3 & & & \\
\vdots & & & \\
-\dfrac{m^N m^1}{\mathcal{M}}I_3 & & & {}_m\!{}^N\left(1-\dfrac{m^N}{\mathcal{M}}\right)I_3
\end{bmatrix}_{3N\times3N}
$$

$$
q = \begin{pmatrix} \alpha^1 \\ \vdots \\ \alpha^N \\ \beta^1 \\ \vdots \\ \beta^N \end{pmatrix}_{6N\times1}, \quad
J = \begin{bmatrix} J^1 & & 0 \\ & \ddots & \\ 0 & & J^N \end{bmatrix}_{3N\times3N}
$$

$$
\ell = \begin{pmatrix} \hat{f} \\ \hat{T} \end{pmatrix}_{6N\times1}, \quad
K = \begin{bmatrix} {}^f_K & 0 \\ 0 & {}^t_K \end{bmatrix}_{6N\times6N}
$$

$$
\tilde{r} = \begin{bmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{bmatrix}_{3\times3}, \quad
\Gamma^T = \begin{bmatrix} {}^f_M\tilde{R} \\ J' \end{bmatrix}_{6N\times3}, \quad
\tilde{R} = \begin{bmatrix} \tilde{R}^1 \\ \vdots \\ \tilde{R}^N \end{bmatrix}_{3N\times3}, \quad
J' = \begin{bmatrix} J^1 \\ \vdots \\ J^N \end{bmatrix}_{3N\times3}
$$

Figure C-2. Flexible Body Appendage Equations

System Equations:

To obtain the system equations write the Newton, Euler equations

$$\underline{F} = M \, \ddot{\underline{R}}$$

$$\underline{T} = \dot{\underline{H}}_v + \dot{\underline{H}}_c \tag{C.23}$$

where

$$\underline{H}_c = \text{momentum from rotors}$$

$$\underline{H}_v \triangleq \int_{\mathcal{M}} \overline{\rho} \times \dot{\overline{\rho}} \, dm \quad \text{vehicle momentum}$$

$$\overline{\rho} = \text{generic position vector from vehicle mass center}$$

Now omitting the details of the expansion of $\underline{H}$ and $\dot{\underline{H}}$ which is contained in Reference 66 we may write immediately after linearizing (128) of Reference[66],

$$T = J^* \dot{\omega} + \dot{H}_c + \tilde{\omega} H_c + \tilde{r} \sum_{i=1}^{N} m^i \, \ddot{\alpha}^i + \sum_{i=1}^{N} (\widetilde{\rho^i + \Delta^i}) m^i \, \ddot{\alpha}^i + \sum_{i=1}^{N} J^i \, \ddot{\beta}^i \tag{C.24}$$

where

$$H_c \triangleq \sum_{j=1}^{k} h_o \, \psi^j + \sum_{i=1}^{\nu} h_i \, \gamma^i \begin{pmatrix} k & \text{CMSs} \\ \nu & \text{reaction wheels} \end{pmatrix} \tag{C.25}$$

$k_i \triangleq$ scalar magnitude of momentum of ith reaction wheel rotor

$h_o \triangleq$ scalar magnitude of momentum of each of k CMG rotors

$\psi^i \triangleq$ unit vector describing direction of momentum vector of ith CMG rotor relative to the frame $\underline{b}$.

$\gamma^j \triangleq$ unit vector in direction of momentum vector of jth reaction wheel rotor relative to frame $\underline{b}$.

$J^* \triangleq$ inertia of vehicle (considered as a rigid body)

Then since $\rho^i + \Delta^i \approx \rho^i$ by assumption of small Δ^i,

$$T = J\dot{\omega} + \sum_{i=1}^{k} h_o \dot{\psi}^i + \sum_{j=1}^{\nu} \dot{h}_j \gamma^j + \sum_{i=1}^{N} \left[m^i (\tilde{r} + \tilde{\rho}^i) \ddot{\alpha}^i \right] + \sum_{i=1}^{N} J^i \ddot{\beta}^i \qquad (C.26)$$

since for CMGs $\dot{h}_o \equiv 0$ and for reaction wheels $\dot{\xi}^i \equiv 0$. Moreover, since ψ^i, i=1-k is a function of the gimbal angles, a matrix S may be defined such that

$$\dot{\psi} = S(\delta)\dot{\delta} \qquad\qquad \dot{\psi} = 3 \times 1 \qquad (C.27)$$

$$S = 3 \times gk$$

$$\delta = gk \times 1$$

where

$$\psi = \psi^1 + \psi^2 + \cdots \psi^k$$

$$\delta^T = \left({}^1\delta_1, \ldots\ {}^1\delta_g,\ {}^2\delta_1, \ldots\ {}^2\delta_g, \ldots\ldots, {}^k\delta_g, \ldots\ {}^k\delta_g \right)$$

for k CMGs, each with g gimbals (g = 1 for single gimbal CMGs and g = 2 for double gimbal CMGs). Then

$$T = J^*\dot{\omega} + h_o\, S(\delta)\dot{\delta} + \sum_{j=1}^{N} \dot{h}_j \gamma^i\, j + \Gamma\, \ddot{q} + \tilde{\omega}\left(h_o\psi(\delta) + \sum_{i=1}^{\nu} h_i(t)\gamma^i \right) \quad (C.28)$$

Now the system equations are

$$J^*\,\dot{\omega} = T - h_o\, {}^cS(\delta)\dot{\delta} - \sum_{i=1}^{\nu} \dot{h}_i(t)\gamma^i - \tilde{\omega}\left(h_o\psi(\delta) + \sum_{i=1}^{\nu} h_i(t)\gamma^i \right) - \Gamma\ddot{q} \quad (C.29)$$

$$M\ddot{q} + Kq = \ell + {}^f\mathcal{M}\ddot{s} - \Gamma^T\dot{\omega} \qquad (C.30)$$

Characterizing the Control Torques, T_c for a Reaction Wheel System.

The control torque may be characterized in many levels of detail, depending upon how deeply one wishes to penetrate into the control hardware dynamics. The level of detail chosen must be compatible with the fidelity of the rest of the model. For linear system

characterizations such as we use here, this implies that the eigen

values of the homogeneous appendage equation (setting right hand side

of (C.30) equal to zero) provide a spectrum within which the "spectrum

of control authority" (those roots associated with the generation of

T_c from measurements which are yet to be specified) must lie. Thus,

erroneous information is obtained from analysis or simulation if we

insist upon modeling modes within the control mechanisms which can

excite structural modes of the physical system but which are absent

from the system model (C.29), (C.30). Likewise, erroneous results are

obtained if there are structural modes modeled in (C.30) which can

interact with modes of the physical controller implementation but

which modes are absent in the model for T_c in (C.29). Therefore, to

be completely honest about the model for T_c we must admit that T_c

cannot be rationally described before an eigenvalue analysis is

completed for the appendage Equations (C.30), and before some controller

hardware is proposed. (In other words we cannot simply write T_c in

(C.29) and proceed with the design or analysis (i.e., optimization)

and then later design the dynamical elements which produce T_c without

returning to check the spectrum compatibilities discussed above).

For the LST mode in (C.29) we have already neglected any shock

mount or elastic suspension of the CMGs or reaction wheels. It is

therefore inappropriate to include any appendage modes or controller

dynamics which exceed the shock mount frequencies. The functional

block diagram of the LST control system is shown in Figure C-3, and the

details of each block are provided in the next section with a view toward

200

completing the system equations which presently have the form in (C.29), (C.30).

Description of the Control System for LST:

For the rate gyro transfer function we have from (see Figure C-3),

$$G_{rg}(s) = \cfrac{1}{\cfrac{s^2}{(2\pi\ 16)^2} + \cfrac{2(.7)}{2\pi\ 16}\,s + 1} = \frac{\omega'(s)}{\theta(s)} \qquad (C.31)$$

or $\ddot{\omega}' + 2\xi_r\omega_r\dot{\omega}' + \omega'\,\omega_r^2 = \omega_r^2\omega_2$, which yields for each root of $G_{rg}(s)$

$$\xi_i = .7$$

$$\omega_r = 2\pi\ 16$$

$$|\lambda_{ry}|_{1,2} = \omega_r = 2\pi\ 16 \qquad (C.32)$$

For the position sensor

$$G_{gs}(s) = \cfrac{1}{\cfrac{s}{2\pi} + 1} = \frac{\theta'}{\theta} \qquad (C.33)$$

or $\dot{\theta}' + \omega_g\theta' = \omega_g\theta$, $\qquad (\omega_g = 2\pi)$

which yields

$$|\lambda_{gs}| = 2\pi \qquad (C.34)$$

For the analog to digital (A/D) prefilter

$$F_p(s) = \cfrac{1}{\cfrac{s}{2\pi\ 20} + 1} = \frac{\omega''}{\omega'} \qquad (C.35)$$

and the associated root

$$|\lambda_p| = 2\pi\ 20 \qquad (C.36)$$

applies to both prefilters for position and rate

$$\dot{\omega}'' + \omega_{rp}\omega'' = \omega_{rp}\omega'$$
$$\dot{\theta}'' + \omega_{rp}\theta'' = \omega_{rp}\theta' \qquad (C.37)$$
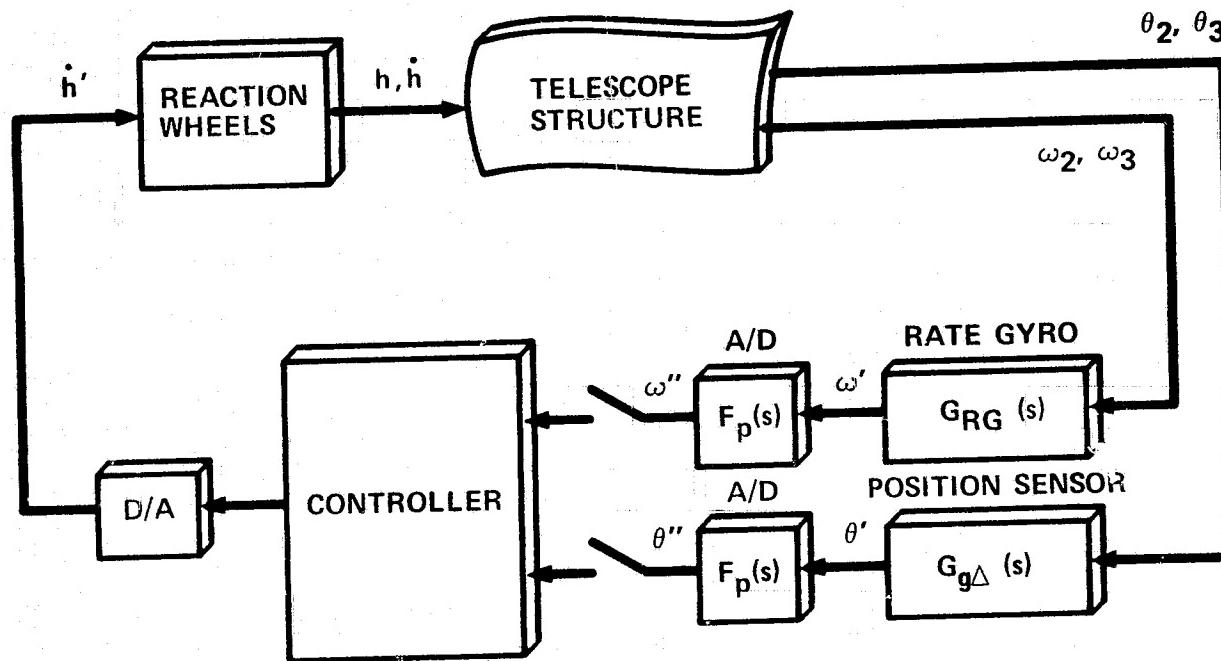
LST BLOCK DIAGRAM

Figure C-3.

The reaction wheel subsystem is illustrated in block diagram form in Figure C-4. Neglecting the friction and the amplifier non-linearities, the governing equations are, for each reaction wheel $j=1,2,3,\ldots \nu$ (let $\nu = 4$).

$$E_1(s) = -\dot{h}'_j - R_F\, i_{F_j}$$

$$E_2(s) = +K_B\, \Omega_j + K_A\, E_1(s)$$

$$i_f(s) = \frac{E_2(s)}{L_m\, s + R_T} \qquad\qquad (C.38)$$

$$\Omega_j = \frac{\dot{h}_j}{J_{R\omega}\, s}$$

$$\dot{h}_j = -K_m\, i_f$$

where h_j is the angular momentum (magnitude) of the jth rotor.

The following is a qualitative argument for studying the linearized equations for the reaction wheel. Some noise (which is always present in the physical circuitry even though it is not modeled yet) will tend to "smooth" a threshold type of nonlinearity as can be illustrated by the use of the dual input describing function (DIDF). Thus, the amplifier threshold and stiction of the wheel are neglected in this model. The elasticity of the bearing/shaft interface allows for linear motion in the small (in the neighborhood of zero wheel speed). For the more frequent non-zero wheel speed circumstance, the friction plays less of a role in the dynamics. Based upon these qualitative arguments, linearity of the reaction wheel model is therefore assumed. Quantitative arguments must await further test data.
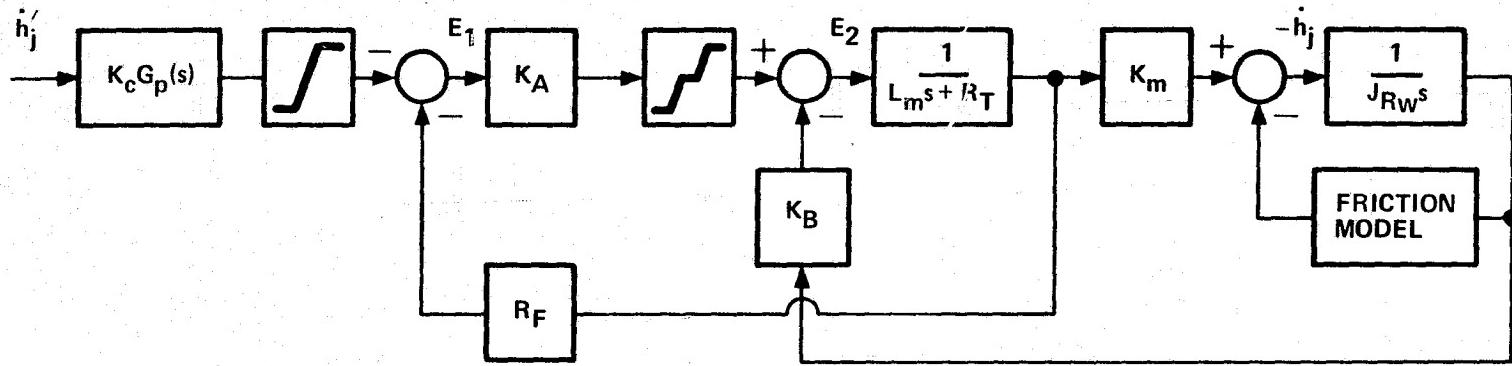
$$K_c G_p(s) = 1$$

$$K_A = 10^4 \text{ v/v}$$

$$L_m = \frac{10}{2\pi 50}$$

$$R_T = 10 \text{ ohms}$$

$$R_F = .316 \text{ to } .48 \text{ ohms}$$

$$K_m = .316 \text{ to } .48 \text{ NM/amp}$$

$$J_{Rw} = .2 \text{ Kg M}^2$$

$$K_B = .316 \text{ to } .48 \text{ v/rad/sec}$$

Figure C-4. Reaction Wheel Control Subsystem.

$\dot{h}'_j$ is the commanded rate of change of the angular momentum of the jth rotor. In differential equation form the reaction wheel model (C.38) becomes

$$\frac{d}{dt}\begin{pmatrix} i_f \\ \Omega \end{pmatrix}_j = \begin{bmatrix} -\left(\dfrac{R_T + K_A R_F}{L_m}\right) & +\dfrac{K_B}{L_m} \\ -\dfrac{K_m}{J_{RW}} & 0 \end{bmatrix} \begin{pmatrix} i_f \\ \Omega \end{pmatrix}_j + \begin{pmatrix} \dfrac{-K_A}{L_m} \\ 0 \end{pmatrix} \dot{h}'_j$$

$$\dot{h}_j = - K_m i_f \qquad\qquad\qquad\qquad (C.39)$$

Or, for the four reaction wheel systems

$$\frac{d}{dt}\begin{pmatrix} i \\ h \end{pmatrix} = \begin{bmatrix} -\left(\dfrac{R_T + K_A R_F}{L_m}\right) I_4 & +\dfrac{K_B}{J_{RW} L_m} I_4 \\ -K_m I_4 & 0 \end{bmatrix} \begin{pmatrix} i \\ h \end{pmatrix} + \begin{bmatrix} -\dfrac{K_A}{L_m} I_4 \\ 0 \end{bmatrix} \dot{h}'$$

$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (C.40)$$

$$h = J_{RW} I_4 \Omega$$

where

$$h = \begin{pmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \end{pmatrix}, \quad i \triangleq \begin{pmatrix} i_{f1} \\ i_{f2} \\ i_{f3} \\ i_{f4} \end{pmatrix} \quad \Omega = \begin{pmatrix} \Omega_1 \\ \Omega_2 \\ \Omega_3 \\ \Omega_4 \end{pmatrix}, \quad \dot{h}' = \begin{pmatrix} \dot{h}'_1 \\ \dot{h}'_2 \\ \dot{h}'_3 \\ \dot{h}'_4 \end{pmatrix}$$

$$I_4 \triangleq 4\times4 \text{ identity matrix}$$

where we have assumed the same hardware parameter values for each of the four reaction wheel subsystems.

The middle term on the right hand side of (C.29) may be written

$$\sum_{i=1}^{4} \gamma^i \dot{h}_i = {}^r S \dot{h} \qquad\qquad\qquad (C.41)$$

205

where

$$r_S = [\gamma^1, \gamma^2, \gamma^3, \gamma^4] \qquad \dot{h}^T \triangleq (\dot{h}_1, \dot{h}_2, \dot{h}_3, \dot{h}_4)$$

Likewise the next to last term in (C.29) may be written

$$\sum_{i=1}^{4} \gamma^i h_i = {}^rS\, h \qquad\qquad\qquad\qquad (C.42)$$

For the rotor mounting arrangement for LST,

$$r_S = \begin{bmatrix} .342 & -.342 & .342 & -.342 \\ .664 & .664 & .664 & .664 \\ .664 & .664 & -.664 & -.664 \end{bmatrix} \qquad (C.43)$$

The system equations, thus far, may be written, neglecting the CMGs in (C.29),

$$J^*\dot{\omega} = T - {}^rS\dot{h} - \tilde{\omega}{}^rSh - \Gamma\ddot{q} \qquad \text{system (rigid body + appendage)}$$

$$M\ddot{q} + D\dot{q} + Kq = \ell + {}^f\!\mathcal{M}\ddot{\Delta} - \Gamma^T\dot{\omega} \qquad \text{appendage}$$

$$\ddot{\omega}_2' + 2\xi_r\omega_r\dot{\omega}_2' + \omega_2'\omega_r^2 = \omega_r^2\omega_2$$

$$\ddot{\omega}_3' + 2\xi_r\omega_r\dot{\omega}_3' + \omega_3'\omega_r^2 = \omega_r^2\omega_3$$

$\left.\right\}$ rate gyros

$$\dot{\theta}_2' + \omega_g\theta_2' = \omega_g\theta_2$$

$$\dot{\theta}_3' + \omega_g\theta_3' = \omega_g\theta_3$$

$\left.\right\}$ position sensors

$$\dot{\omega}_2'' + \omega_{rp}\omega_2'' = \omega_{rp}\omega_2'$$

$$\dot{\omega}_3'' + \omega_{rp}\omega_3'' = \omega_{rp}\omega_3'$$

$$\dot{\theta}_2'' + \omega_{gp}\theta_2'' = \omega_{gp}\theta_2'$$

$$\dot{\theta}_3'' + \omega_{gp}\theta_3'' = \omega_{gp}\theta_3'$$

$\left.\right\}$ prefilters, F_p

$$\dot{h} = -K_m i$$

$$\dot{i} = -p_1 i + p_2 h - p_3 \dot{h}'$$

$\left.\right\}$ reaction wheels

The collection of these equations is numbered (C.44).

External Disturbances:

The external disturbances acting on LST, in approximate order of their importance, include:

1. gravity gradient

2. aerodynamic

3. magnetic

4. thermal

5. solar pressure

6. micrometeorites

Those internal effects which have not been included in the model thus far include disturbing effects from:

1. rotating machinery (vibrations from unbalanced rotors)

2. sensor and electronic noise

3. equipment motion (pumps, shutters, etc.)

4. instrumentation anomalies (bias errors in instrument

   alignment and amplifiers)

The first four of the external disturbance effects listed above are considered important for LST. The thermal effects (4) have been considered in the derivation of the flexible spacecraft model (C.44). The first three, therefore, remain to be described.

Gravity Gradient Torques:

Derivations of gravity gradient torques are found in several sources, but the expressions in Reference 67 will be most useful here. From Equation (B.22) in Reference 67, for circular orbit,

$$\underline{T}^g = 3\Omega_o^2 \, \underline{\xi}_3 \times \underline{\underline{J}} \cdot \underline{\xi}_3 \tag{C.45}$$

where

$$\underline{\xi} \triangleq N \underline{z} \quad , \quad \underline{b} = \Theta \underline{z} \quad , \quad \underline{\xi} = N \Theta^T \underline{b}$$
$$\Theta = I - \widetilde{\theta}$$

$$N = \begin{bmatrix} cn_t & -sn_x sn_t & -cn_x sn_t \\ 0 & cn_x & -sn_x \\ sn_t & sn_x cn_t & cn_x cn_t \end{bmatrix} \qquad\qquad\qquad (C.46)$$

$$\widetilde{\theta} = \begin{bmatrix} 0 & -\theta_3 & \theta_2 \\ \theta_3 & 0 & -\theta_1 \\ -\theta_2 & \theta_1 & 0 \end{bmatrix}$$

n_x = angle between orbital plane and sum line

n_t = vehicle orbital position measured from orbital midnight

$\quad$ ($n_t = \Omega_o t$ for circular orbit)

Ω_o = orbital rate

$\mathbb{J}$ = inertia dyadic of vehicle (assume rigid)

$$\{\underline{b}\}^T \ J \ \{\underline{b}\}$$

We wish to obtain an expression for the torques in $\underline{b}$ coordinates and

linearized about $\theta = 0$.

$$\underline{T}^g = \{\underline{b}\}^T \ T^g$$
$$T^g \approx T_o^g + G^g \ \theta \qquad\qquad\qquad (C.47)$$

To proceed, define from (C.46)

$$n_3^T \triangleq (sn_t, \ sn_x \ cn_t, \ cn_x \ cn_t)$$
$$q^T \triangleq n_3^T \ \Theta^T = n_3^T [I + \widetilde{\theta}] \qquad\qquad\qquad (C.48)$$

Then (C.45) may be written in $\underline{b}$ coordinates as

$$T^g = 3\Omega_o^2 \ \widetilde{q} \ J \ q \qquad\qquad\qquad (C.49)$$

208

But from (C.48), (C.49) becomes

$$T^g = 3\Omega_o^2 \; ([I-\tilde{\theta}]n_3)^{\sim} \; J[I-\tilde{\theta}]n_3$$

$$\approx 3\Omega_o^2 \left\{ \tilde{n}_3 \; J \; n_3 + \tilde{n}_3 \; J \; \tilde{n}_3 \; \theta - \left(\tilde{\theta} \; n_3 \right)^{\sim} J \; n_3 \right\} \tag{C.50}$$

discarding second order terms in θ. A useful identity

$$(\tilde{\theta}n_3)^{\sim} = n_3 \; \theta^T - \theta \; n_3^T$$

gives (C.50) the form

$$T^g = 3\Omega_o^2 \left\{ \tilde{n}_3 \; Jn_3 + \tilde{n}_3 \; J \; \tilde{n}_3 \; \theta - n_3 n_3^T \; J\theta + n_3^T \; Jn_3\theta \right\} \tag{C.51}$$

Hence, comparing (C.51) to (C.47)

$$\left. \begin{array}{l} T_o^g = 3\Omega_o^2 \; \tilde{n}_3 \; J \; n_3 \\[2mm] G^g = 3\Omega_o^2 \left\{ \tilde{n}_3 \; J \; \tilde{n}_3 - n_3 n_3^T \; J + n_3^T \; J \; n_3 \; I \right\} \end{array} \right\} \tag{C.52}$$

For the case

$$n_x = 45° \qquad , \qquad \Omega_o = .00111 \; rad/s$$

$$J = \begin{bmatrix} 18,357 & 0 & 0 \\ 0 & 41,822 & 0 \\ 0 & 0 & 44,484 \end{bmatrix} \; Kg \; m^2$$

T_o^g in (C.52) becomes

$$T_o^g = 3(.00111)^2 \begin{pmatrix} n_{3_2} \; n_{3_3} \; (J_{33} - J_{22}) \\ n_{3_1} \; n_{3_3} \; (J_{11} - J_{33}) \\ n_{3_1} \; n_{3_2} \; (J_{22} - J_{11}) \end{pmatrix} = 3\Omega_o^2 \begin{pmatrix} \Delta_{32} \; sn_x cn_x \; c^2\Omega_o t \\ \Delta_{13} \; cn_x \; \frac{1}{2} s2\Omega_o t \\ \Delta_{21} \; sn_x \; \frac{1}{2} s2\Omega_o t \end{pmatrix}$$

$$= 3\Omega_o^2 \begin{pmatrix} \frac{1}{2}(1+ c2\Omega_o t) \; s \; n_x cn_x \; \Delta_{32} \\ \frac{1}{2} s2\Omega_o t \quad cn_x \quad \Delta_{13} \\ \frac{1}{2} s2\Omega_o t \quad sn_x \quad \Delta_{21} \end{pmatrix}$$

$$= \frac{3\Omega_o^2}{2} \begin{bmatrix} sn_x cn_x \Delta_{32} & 0 & sn_x cn_x \Delta_{32} \\ 0 & cn_x \Delta_{13} & 0 \\ 0 & sn_x \Delta_{21} & 0 \end{bmatrix} \begin{pmatrix} 1 \\ s \; 2\Omega_o t \\ c \; 2\Omega_o t \end{pmatrix}$$

Using the above data,

$$T_o^g = \begin{bmatrix} .002 & 0 & .002 \\ 0 & -.034 & 0 \\ 0 & .031 & 0 \end{bmatrix} \begin{pmatrix} 1 \\ s \; 2\Omega_o t \\ c \; 2\Omega_o t \end{pmatrix} \quad \text{n.m.} \qquad (C.53)$$

$$G^g = \begin{bmatrix} 0 & .0035s2n_t & -.0035s2n_t \\ .034s2n_t & -.089 + .059c2n_t & -.024 - .024 c2n_t \\ .031s2n_t & -.022 - .022c2n_t & -.022 - .065 c2n_t \end{bmatrix} \qquad (C.54)$$

where for circular orbit $n_t = \Omega_o t$. Note that due to the small angles θ_i the term T_o^g in (C.47) is more significant in magnitude than is the $G^g \theta$ term.

Magnetic Torques:

An approximation of the geomagnetic potential, taken from Equation (6) of Reference 68, is

$$v \approx r_o \sum_{n=1}^{\infty} \sum_{m=0}^{n} \left(\frac{r_o}{R}\right)^{n+1} (g_n^m \cos m\phi + h_n^m \sin m\phi) P_n^m(\lambda-) \qquad \text{(C.55)}$$

where from Figure 4, p. 17 of Reference 68,

r_o = radius of earth

R = radius of orbit

$\bar{\lambda}$ = $1/2 - \lambda$

λ = latitude

ϕ = easterly longitude

g_n^m , h_n^m = Schmidt coefficients (see Table 1 in Reference 68)

$P_n^m(\bar{\lambda})$ = Legendre functions with Schmidt normalization

The magnetic field in a region containing no sources satisfies

$$\underline{B} = - \underline{\nabla} V \qquad \text{(C.56)}$$

and the torque applied to the spacecraft is

$$\underline{T}^m = \underline{M} \times \underline{B} \qquad \underline{B} = \{\underline{b}\}^T B_b$$

$$\underline{M} = \{\underline{b}\}^T M_b \qquad \text{(C.57)}$$

where $\underline{M}$ is the magnetic field generated within the spacecraft. Now we wish to truncate the series for v and express $\underline{T}^m$ in the vehicle coordinate frame, $\underline{b}$. To relate λ and ϕ to the n_x and n_t of previous sections define the coordinate frame $\underline{n}$ as in Figure C-5.

The relation between the $\underline{\xi}$ frame and the spherical coordinate frame $\underline{s}$ in the figure is

$$\underline{s} = \begin{bmatrix} s_\beta & c_\beta & 0 \\ -c_\beta & s_\beta & 0 \\ 0 & 0 & 1 \end{bmatrix} \underline{\xi} = \bar{B} \, \underline{\xi} \qquad \text{(C.58)}$$
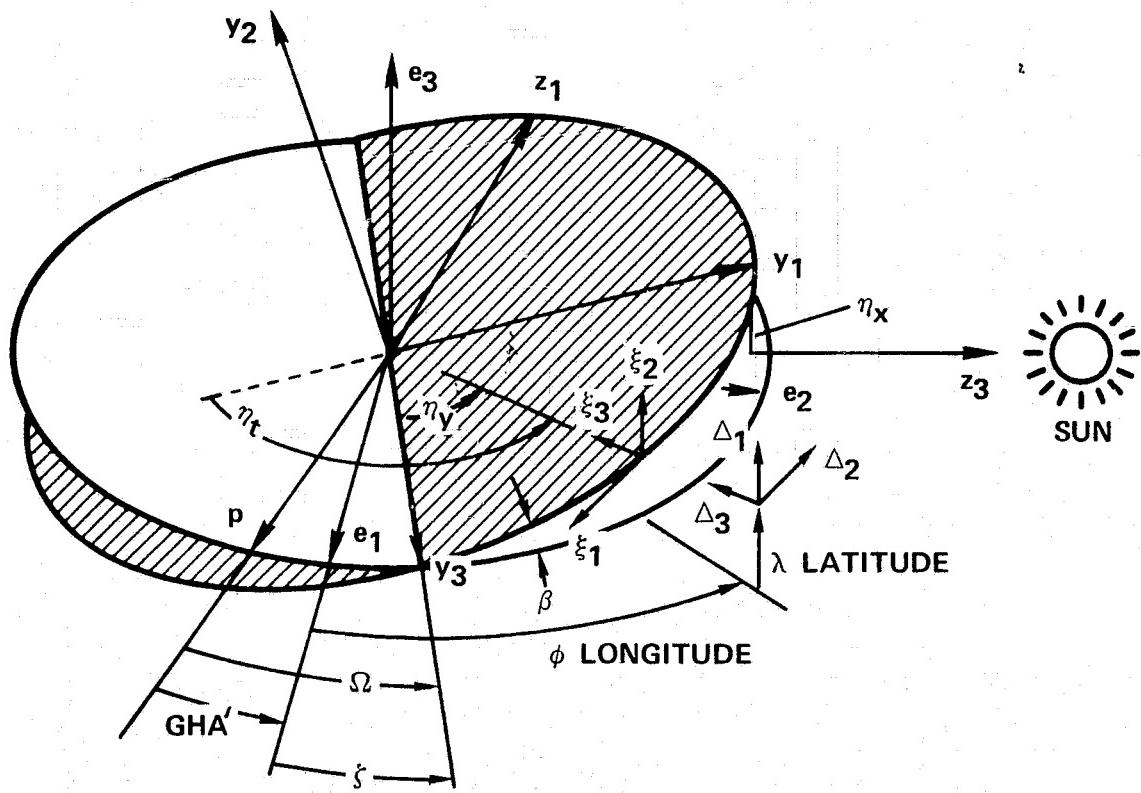
Figure C-5. Coordinates for Magnetic Torques.

Now the magnetic field in $\underline{s}$ coordinates is, from Equation (8) in Reference 68, with changes in notation,

$$\begin{pmatrix} B_1 \\ B_2 \\ B_3 \end{pmatrix}_s = \begin{pmatrix} \dfrac{1}{R}\dfrac{\partial v}{\partial \bar{\lambda}} \\ -\dfrac{1}{R\sin\bar{\lambda}}\dfrac{\partial v}{\partial \phi} \\ \dfrac{\partial v}{\partial R} \end{pmatrix} \tag{C.59}$$

where

$$\bar{\lambda} \triangleq \pi/2 - \lambda$$

$$\underline{B} = \{\underline{s}\}^T B_s = \{\underline{b}\}^T B_b$$

Now since from Reference 67, p. A-5

$$\underline{\xi} = Q\,\underline{b} \quad , \tag{C.60}$$

where

$$Q = \begin{bmatrix} Ct & -sxst & -cxst \\ 0 & cx & -sx \\ st & sxct & cxct \end{bmatrix} [I + \tilde{\theta}]$$

From (C.58), (C.60),

$$\underline{s} = \bar{B}\,Q\,\underline{b} \tag{C.61}$$

$$\bar{B}Q = \underbrace{\begin{bmatrix} s\beta ct & -s\beta sxst + c\beta cx & -s\beta cxst - c\beta sx \\ -c\beta ct & c\beta sxst + s\beta cx & c\beta cxst - s\beta sx \\ st & sxct & cxct \end{bmatrix}}_{\bar{Q}} [I + \tilde{\theta}]$$

Thus

$$B_s = \bar{B}\,Q\,B_b$$

$$B_b = Q^T\,\bar{B}^T\,B_s \tag{C.62}$$

213

Since from Equation (11) of Reference 68,

$$B_s = -g_1^o \left(\frac{r_o}{R}\right)^3 \begin{pmatrix} \sin\overline{\lambda} \\ 0 \\ 2\cos\overline{\lambda} \end{pmatrix}$$

for a simple dipole model of the magnetic field (only the first term of the spherical harmonic expansion in (C.55) is needed). From Reference 69.

$$g_1^o = -30425 \quad \text{gauss}$$

$$r_o = 6371.2 \times 10^3 \text{ m}$$

The angle $\overline{\lambda}$ can be related to n_t from the geometry of Figure C-5 using spherical trigonometry.

$$\sin(n_y - \pi + n_t) = \frac{\sin\lambda}{\sin\beta} = \frac{\sin(\pi/2 - \overline{\lambda})}{\sin\beta} = \frac{\cos\overline{\lambda}}{\sin\beta}$$

Then by the sum of angles indentity

$$\cos\overline{\lambda} = \sin\beta(\sin(n_y - \pi)\cos n_t + \cos(n_y - \pi)\sin n_t)$$

$$\sin\overline{\lambda} = \sqrt{1 - \cos^2\overline{\lambda}} \qquad (\text{since } 0 \leq \overline{\lambda} \leq \pi)$$

Then

$$B_s = -g_1^o\left(\frac{r_o}{R}\right)^3 \begin{pmatrix} \left[1 - s^2\beta(S_{y-\pi}^2 c_t^2 + c_{1-\pi}^2 s_t^2 + 2S_{y-\pi}c_{y-\pi}S_t c_t)\right]^{1/2} \\ 0 \\ 2.S\beta(S(n_y - \pi)Ct + C(n_y - \pi)St) \end{pmatrix} \qquad (C.63)$$

and from (C.62)

$$B_b = -g_1^o \left(\frac{r_o}{R}\right)^3 (\overline{B}\ Q)^T \left(\begin{array}{c} [1-S^2\beta(S^2(n_y-\pi)C_t^2+C^2(n_y-\pi)S_t^2+2S(n_y-\pi)c(n_y-\pi)S_tC_t)]^{\frac{1}{2}} \\ 0 \\ 2S\beta(s(n_u-\pi)c_t + c(n_y-\pi)St) \end{array}\right)$$

$$\underbrace{}_{\lambda'}$$

$$(C.64)$$

By inspection of (C.57) and (C.64) it is apparent that B_b is not simply sinusoidal in time and that the linearized form of the magnetic torques

$$T^m = \widetilde{M}_b\ B_b \approx T_o^m + G^m\theta$$

will yield terms in T_o^m and G^m which will, if expanded in a Fourier series, contain higher harmonics of the orbital frequency. Specifically it can be shown from (C.61), (C.64) that

$$T_o^m = -\ q_1^o \left(\frac{r_o}{R}\right)^3 \overline{Q}^T\ \lambda' \qquad\qquad (C.65)$$

and

$$G^m = -\ g_1^o \left(\frac{r_o}{R}\right)^3 \left(\widetilde{\overline{Q}^T\lambda'}\right) = \widetilde{T}_o^m \qquad\qquad (C.66)$$

Equation (C.65) could now be expanded into a Fourier series in time in order to obtain the coefficients of the various harmonics of orbital frequency. Of specific interest will be the first three or four harmonics

$$T_o^m = T_{oo}^m + T_{01s}^m \sin n_t + T_{01c}^m \cos n_t + T_{02s}^m \sin 2n_t \qquad (C.67)$$

$$+ T_{02c}^m \cos 2 n_t + T_{03s}^m \sin 3 n_t + T_{03c}^m \cos 3 n_t +---$$

215

The magnitude of the magnetic field assumed for the LST spacecraft is 100 Gauss. The first four harmonics of $T_o^m(t)$ have been approximated for use in the text as

$$T_o^m = .003(\sin .00111t + \sin .00222t + \sin .00333t + \sin .00444t)$$
(C.68)

Finally, the linearized disturbances from gravity gradient and magnetic torques can be written (aerodynamic torques have been considered only to add some small amplitude to the Fourier terms comprising gravity gradient and magnetic torques.)

$$T = (T_o^q + T_o^m) + [G^g + G^m]\theta = T_o + G\theta$$
(C.69)

so that the total system equations may now be written, substituting $\dot{\theta}$ for ω in the linearized equations

$J^* \ddot{\theta} + \Gamma \ddot{q} = -\,^rS\dot{h} + (^rSh)^\sim \dot{\theta} + G\theta + T_o(t)$ }rigid body + appendage

$M\ddot{q} + Kq + \Gamma^T \ddot{\theta} = \,_f^f\mathcal{M}\ddot{\Delta}(t)$ }appendage

$\dot{h} = -K_m i$

$\dot{i} = -p_1 i + p_2 h - p_3 \dot{h}'$ }reaction wheels

$\ddot{\omega}_2' + 2\xi_r \omega_r \dot{\omega}_2' + \omega_r^2 \omega_2^1 = \omega_r^2 \dot{\theta}_2$

$\ddot{\omega}_3' + 2\xi_r \omega_r \dot{\omega}_3' + \omega_r^2 \omega_3' = \omega_r^2 \dot{\theta}_3$ }rate gyro

(C.70)

$\dot{\theta}_2' + \omega_g \theta_2' = \omega_g \theta_2$

$\dot{\Theta}_3' + \omega_g \Theta_3' = \omega_g \theta_3$ }position sensor

$\dot{\omega}_2'' + \omega_{rp} \omega_2'' = \omega_{rp} \omega_2'$

$\dot{\omega}_3'' + \omega_{rp} \omega_3'' = \omega_{rp} \omega_3'$ }Prefilter rate signal

$\dot{\theta}_2'' + \omega_{gp} \theta_2'' = \omega_{gp} \theta_2'$

$\dot{\theta}_3'' + \omega_{gp} \theta_3'' = \omega_{gp} \theta_3'$ }prefilter, position signal

where external forces and torques on the appendage are neglected ($\ell = 0$).

To proceed to study this system of equations it is helpful to first make some coordinate transformations to further simplify its structure. Even if the equations were time invariant, an eigenvalue analysis of these original equations might be too large a task for the computer. Thus, two steps may be taken to simplify the task. First, replace the time varying G with its value averaged over one orbit. (It can be seen in the derivations of the external disturbances that the Gθ term is smaller than the T_o term in (C.69) if θ is small). Also, long experiments (5 hours) are required by LST and the averaging of G over one orbit (≈ 1.5 hours) seems acceptable. $T_o(t)$, however is not to be averaged, since its time dependence may be important to the control problem.

A further simplification is available if the cross product between vehicle rate and rotor momentum is neglected. In this event the term $(^r\bar{Sh})^{\sim}\dot{\theta}$ in (C.70) is discarded. This step is more justified if the nominal momentum is zero.

The appendage equations can be transformed to distributed or "modal" coordinates, η , by the choice

$$q = \Phi\eta \qquad\qquad\qquad\qquad (C.71)$$

where Φ is the matrix of assumed independent eigenvectors satisfying the orthonormality relationships (see Reference 66 for further discussion).

$$\phi^{k^T} M \phi^k = 1 \qquad k=1,\ldots,\nu$$

$$\phi^{k^T} M \phi^j = 0 \qquad k \neq j$$

$$\phi^{k^T} M \phi^j = 0 \qquad k \neq j$$

$$\Phi = [\phi^1 \quad \phi^2 \; \text{---} \; \phi^\nu]$$

Then the appendage equation takes on the form, using (C.71) and pre-multiplying (C.30) by Φ^T,

$$\ddot{\eta} + 2\zeta\,\sigma\,\dot{\eta} + \sigma^2\,\eta + \Phi^T\Gamma^T\ddot{\Theta} = \Phi^T f_m \ddot{\Delta}(t) \qquad \text{(C.72)}$$

where modal damping ζ, is added in this form.

Now the first two equations of (C.70) may be written

$$
\begin{bmatrix} J^* & \Gamma\Phi \\ (\Gamma\Phi)^T & I_\nu \end{bmatrix}
\begin{pmatrix} \ddot{\theta} \\ \ddot{\eta} \end{pmatrix}
=
\begin{bmatrix} {}^r\widetilde{Sh} & 0 \\ 0 & -2\zeta\sigma \end{bmatrix}
\begin{pmatrix} \dot{\theta} \\ \dot{\eta} \end{pmatrix}
+
\begin{bmatrix} G & 0 \\ 0 & -\sigma^2 \end{bmatrix}
\begin{pmatrix} \theta \\ \eta \end{pmatrix}
$$
$$
+
\begin{bmatrix} {}^r S\, K_m \\ 0 \end{bmatrix} i
+
\begin{bmatrix} I_3 & 0 \\ 0 & (\Phi^T)(f_{\mathcal{M}}) \end{bmatrix}
\begin{pmatrix} T_o \\ \ddot{\Delta} \end{pmatrix}
\qquad \text{(C.73)}
$$

Solving for the highest derivatives

$$
\begin{pmatrix} \ddot{\theta} \\ \ddot{\eta} \end{pmatrix}
= L_1
\begin{pmatrix} \dot{\theta} \\ \dot{\eta} \end{pmatrix}
+ L_2
\begin{pmatrix} \theta \\ \eta \end{pmatrix}
+ L_3\, i + L_4
\begin{pmatrix} T_o \\ \ddot{\Delta} \end{pmatrix}
\qquad \text{(C.74)}
$$

where

$$
L_1 \triangleq \hat{J}^{-1}
\begin{bmatrix} {}^r\widetilde{Sh} & 0 \\ 0 & -2\zeta\sigma \end{bmatrix}
, \qquad
L_2 \triangleq \hat{J}^{-1}
\begin{bmatrix} G & 0 \\ 0 & -\sigma^2 \end{bmatrix}
$$

$$
L_3 \triangleq \hat{J}^{-1}
\begin{bmatrix} {}^r S K_m \\ 0 \end{bmatrix}
, \qquad
L_4 \triangleq \hat{J}^{-1}
\begin{bmatrix} I_3 & 0 \\ 0 & \Phi^T f_{\mathcal{M}} \end{bmatrix}
$$

$$
\hat{J} \triangleq
\begin{bmatrix} J^* & \Gamma\,\Phi \\ \Phi^T\,\Gamma^T & I_\nu \end{bmatrix}
, \quad
\hat{J}^{-1} =
\begin{bmatrix} F^{-1} & -F^{-1}\,\Gamma\,\Phi \\ -\Phi^T\,\Gamma^T F^{-1} & I + \Phi^T\Gamma^T F^{-1}\Gamma\,\Phi \end{bmatrix}
\qquad \text{(C.75)}
$$

$$F \triangleq J^* - \Gamma \, \Phi \, \Phi^T \, \Gamma^T$$

Now define the state vector x^1, the control vector u^o, the output vector, y^1, and the measurement vector, z^1.

$$
x^1 \triangleq
\begin{bmatrix}
\theta''_2 \\
\theta''_3 \\
\hline
\theta'_2 \\
\theta'_3 \\
\hline
\omega''_2 \\
\omega''_3 \\
\hline
\omega'_2 \\
\omega'_3 \\
\hline
\dot{\omega}'_2 \\
\dot{\omega}'_3 \\
\hline
\theta \\
\eta \\
\hline
\dot{\theta} \\
\hline
\dot{\eta} \\
\hline
h_\Delta \\
i
\end{bmatrix}
=
\begin{bmatrix}
23_{\theta}{}'' \\
\hline
23_{\theta}{}' \\
\hline
23_{\omega}{}'' \\
\hline
23_{\omega}{}' \\
\hline
23_{\dot{\omega}}{}' \\
\hline
\theta \\
\eta \\
\hline
\dot{\theta} \\
\hline
\dot{\eta} \\
\hline
h_\Delta \\
i
\end{bmatrix}
, \quad u^o \triangleq \dot{h}
$$

$$y^{1^T} \triangleq (\theta_2, \ \theta_3, \ \dot{\theta}_2, \ \dot{\theta}_3)$$

$$z^{1^T} \triangleq (\theta''_2, \ \theta''_3, \ \omega''_2, \ \omega''_3)$$

$$\text{(C.76)}$$

where

$$h_\Delta \triangleq h - \overline{h}$$

$\overline{h}$ = current value of h(t) at time of update of linearized model

Then the linear state equations may be written

$$\dot{x}^1 = A^1\,x^1 + B^1\,u^o + \Sigma^1\,w^1$$
$$y^1 = C^1\,x^1 \qquad\qquad (C.77)$$
$$z^1 = M^1\,x^1$$

where

$$
A^1 \triangleq
\begin{bmatrix}
-\omega_{gp}I_2 & \omega_{gp}I_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -\omega_g I_2 & 0 & 0 & 0 & 0 & \omega_g I_2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -\omega_{rp}I_2 & \omega_{rp}I_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & I_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -\omega_r^2 I_2 & -2\xi_r\omega_r I_2 & 0 & 0 & 0 & 0 & \omega_r^2 I_2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & & I_{3+\nu} & & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & & L_2 & L_1 & & 0 & L_3 \\
0 & 0 & 0 & 0 & 0 & 0 & & 0 & & 0 & -K_m I_4 \\
0 & 0 & 0 & 0 & 0 & 0 & & 0 & p_2 I_4 & -p_1 I_4
\end{bmatrix}
$$

with column labels $23_\theta''$, $23_\theta'$, $23_\omega''$, $23_\omega'$, $23_{\dot\omega}'$, $\theta_1\,23_\theta\,\eta$, $\dot\theta_1\,23_{\dot\theta}\,\dot\eta$, h_Δ, i and row labels $23_\theta''$, $23_\theta'$, $23_\omega''$, $23_\omega'$, $23_{\dot\omega}'$, $\theta\,\eta$, $\dot\theta\,\dot\eta$, h_Δ, i.

$$\qquad (C.78)$$

$$
B^1 \triangleq
\begin{bmatrix}
0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -p_3 I_4
\end{bmatrix},
\qquad
\Sigma^1 \triangleq
\begin{bmatrix}
0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ L_4 & 0 \\ 0 & 0 \\ 0 & p_2
\end{bmatrix},
\qquad
w^1 =
\left(
\frac{\begin{pmatrix} T_o \\ \ddot{\Delta} \end{pmatrix}}{\overline{h}}
\right)
\qquad (C.79)
$$

$$
\begin{array}{c}
\quad\;\; \theta'' \quad \theta' \quad \omega'' \quad \omega' \quad \dot{\omega}' \quad \theta_1 \quad \theta_2 \quad \theta_3 \quad \eta \quad \dot{\theta}_1 \quad \dot{\theta}_2 \quad \dot{\theta}_3 \quad \dot{\eta} \;\; h_\Delta \;\; i
\end{array}
$$

$$
C^1 = \left[\begin{array}{ccccc|ccc|c|ccc|c|c}
 &&&&& 0 & 1 & 0 && 0 & 0 & 0 && \\
 &&&&& 0 & 0 & 1 && 0 & 0 & 0 && \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
 &&&&& 0 & 0 & 0 && 0 & 0 & 1 && \\
\end{array}\right]
$$

$$
M^1 = \left[\begin{array}{cc|c|cc|c|c|c|c}
1 & 0 && 0 & 0 &&&& \\
0 & 1 && 0 & 0 &&&& \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 && 0 & 1 &&&& \\
\end{array}\right]
$$

$$
\text{(C.80)}
$$

Steps are outlined in section 6.0 of the text to reduce such a model for control design.